Annex to:

EFSA NDA Panel, 2022. Scientific Opinion on the Tolerable Upper Intake Level for dietary sugars. EFSA Journal 2022;20(2):7074. doi:10.2903/j.efsa.2022.7074

Annex L – Statistical analyses performed on intervention studies retrieved from published scientific literature as preparatory work for the setting of a Tolerable Upper Intake Level for dietary sugars

European Food Safety Authority (EFSA)

Summary

The objective of this statistical analysis was to investigate the effect of added and free sugars on the risk of metabolic diseases in adults and children using evidence from intervention studies. When appropriate the relationship was also characterised with meta-regressive dose–response analysis. Since incidence of disease was not measured in any of the experimental studies due to the short duration of the intervention, surrogate endpoints were considered instead as indicators of metabolic disease risk (e.g. fasting glucose levels for type 2 diabetes). The analysis was based on the data collected via a systematic review.

The steps below were followed in the statistical analysis of the experimental studies:

1) For each endpoint (e.g. body weight, fasting glucose levels), the studies were clustered according to the subquestions they could answer: Q1 refers to the effect of the amount of sugars and Q2 to the effect of the type of sugar.
2) Forest plots were produced to describe visually the relationships between sugars intake, expressed as percentage of energy intake (E%), and all the endpoints considered at the level of individual studies. In the graphical display, the pooled mean effect was reported, along with its 95% confidence and prediction intervals (95% confidence interval (CI) and 95% PI).
3) A mixed-effects meta-regressive dose–response model was set up to characterise and quantify the relationship between the intake of sugars (E%) and a subset of metabolic endpoints (i.e. fasting triglycerides, fasting glucose, fasting insulin, body weight and uric acid) in the population subgroups investigated in the body of evidence (BoE). Both the linear and non-linear shapes were explored.
4) Each model was adjusted for a set of explanatory factors (fixed effects) and a set of factors explaining the hierarchical structure of the data (random effects).

These steps are described in detail in the following sections.

Table of Contents

Subquestions to address and overall approach taken

The analyses conducted on the intervention studies aimed at answering the following two subquestions:

Q1: Is the intake of (total/added/free) sugars positively and causally associated with the risk of chronic metabolic diseases at the levels of intake and in the population subgroups investigated in the studies eligible for this assessment? For this question randomised controlled trials (RCTs) providing arms with different amounts of sugar (e.g. fructose; mixtures of fructose and glucose as sucrose, sugar-sweetened beverages (SSBs), honey; non-milk extrinsic sugars, simple carbohydrates from the whole diet) were selected. Of these, four studies targeted free sugars and the rest only manipulated the added sugars fraction. Since the intakes of added and free sugars widely overlap, RCTs addressing Q1 were combined to draw conclusions on added and free sugars, even if the majority manipulated only the added sugars fraction. These studies, however, did not allow conclusions on total sugars.

Q2: Is the intake of fructose positively and causally associated with the risk of chronic metabolic diseases at the levels of intake and in the population subgroups investigated in the studies eligible for this assessment? For this question RCTs comparing the same amount of fructose and glucose were considered.

Evidence synthesis

1.1. Data pre-processing

In this step, the data were edited and standardised to be prepared for the statistical analysis.

First, all the measurement units were standardised. Subsequently the mean effect for each study was calculated.

1.1.1. Data standardisation

The sugars intake was converted to per cent energy (E%) when reported in amount consumed per day (g/day) following the formula based on the average energy content of sugars (4 kcal/g):

$$\%E = \frac{4*(g/day)}{x}, \quad \textbf{Equation 1}$$

where *x* represents the study mean daily energy intake. If *x* was not reported in the publication, it was been imputed using standard values as follows:

- x = 2,000 kcal if the study population was mixed males and females;
- x = 1,800 kcal if the study population was only females;
- x = 2,200 kcal if the study population was only males.

The variation parameter (*VP*) was always converted into a standard error (SE) using to the formulas below:

1) if the *VP* was expressed as 95% CI, assuming the data are normally distributed:

$$\boldsymbol{SE = \frac{UL - LL}{3.92}} \quad \textbf{Equation 2}$$

where UL and LL are respectively the upper and lower boundary of the CI.

2) if the *VP* was expressed as interquartile range, assuming the data are normally distributed:

$$\boldsymbol{SE = \frac{IQUL - IQLL}{3.5}} \quad \textbf{Equation 3}$$

where IQUL and IQLL are respectively the upper and lower boundary of the interquartile range.

3) if the *VP* is expressed as absolute range, assuming the data are normally distributed:

$$SE = \frac{NRUL - NRLL}{4}$$ **Equation 4**

where NRUL and NRLL are respectively the upper and lower boundary of the range.

4) if *VP* is expressed an absolute range, assuming the data are not normally distributed:

$$SE = \frac{RUL - RLL}{6}$$ **Equation 5**

where RUL and RLL are respectively the upper and lower boundary of the range.

5) if the *VP* is expressed as standard deviation,

$$SE = \frac{SD}{\sqrt{n}}$$ **Equation 6**

where *n* is the sample size.

Subsequently, the location parameter (*LP*) was always converted into a mean as follows:

6) if the LP was expressed as a Median

$$\mu := Mean = \frac{RUB + RLB + 2Median}{4}$$ **Equation 7**

where RUB and RLB are respectively the upper and lower boundary of the range.

Particular attention was paid to the mean values at the end of the intervention (T1) when expressed as function of the mean at the beginning of the intervention (T0). These additional transformations were computed as follows:

7) if μ at time T1 (μ_{T1}) is expressed as a rate of change of the mean at T0 *(Δ)*

$$\mu_{T1} = \mu_{T0}\,(1 + \Delta)$$ **Equation 8**

8) if μ_{T1} is expressed as absolute change of the mean at T0 *(AΔ)*

$$\mu_{T1} = \mu_{T0} + A\Delta$$ **Equation 9**

9) if $\mu_{T1}$is expressed as a percentage value of the mean value at the initial time *(%value)*

$$\mu_{T1} = \mu_{T0} \cdot \%value$$ **Equation 10**

SE values were transformed accordingly when needed.

Finally, the measurement unit of each endpoint were standardised and, when needed, converted following the respective formulas reported in 0. For endpoints related to ectopic fat deposition such as liver fat and visceral adipose tissue (VAT), the standardisation of the measurement unit was not possible.

1.1.2. Mean effect computation

A decision was made to consider the difference in sugars intake between arms as the variable of interest, rather than the sugars dose administered to each intervention group. This choice was led by the nature of the intervention that, for most studies, did not address the whole diet but only part of it. Therefore, only the amount of sugars consumed with that fraction of the diet was reported. Only rarely the sugars intake from the background diet could be estimated. This circumstance prevented the possibility to use in the analyses the intake of sugars as reported in the studies, being the latter a potentially inaccurate estimate of the true intake (a sugars intake of 10 E% provided with the intervention could correspond, for instance, to a sugars intake of 20 E% from all sources considering also the fraction of the diet not controlled by the intervention), and led to the computation of the difference in sugars intake between arms. Consequently, the mean level of the endpoint achieved at the end of the intervention in each treatment group could not be used, and a measure of the treatment effect had to be calculated. Due to the continuous nature of the endpoints, a natural choice for the effect measure was to use a difference of the endpoint mean at the end of the treatment between arms, or a difference of the mean change from baseline between arms.

The effect measure computation required identifying the control and intervention arms, better fitting questions Q1 and Q2, from each study. This was carried out according to the procedure below:

- Q1 assesses the effect of the amount of sugars and comparisons are made between:
  - one arm with a zero (added or free) sugars dose (control) and one arm with sugars dose > 0 (intervention) that could be any type of sugar;
  - two arms with different doses of the same sugar in which the arm with the lowest dose is the control and the arm with the highest dose the intervention.
- Q2 investigates the effect of the type of sugar, and comparisons are made between arms that provide the same amount of glucose (control) and fructose (intervention).

The study arms selected to address the previous questions are reported in Appendix B –.

The sugars dose (d) was calculated as the difference in sugars intake between the intervention and control arms:

$$\boldsymbol{d = d_{int} - d_{contr}} \qquad \textbf{Equation 11}$$

The scheme of the mean effect calculation procedure is summarised in Figure 1.

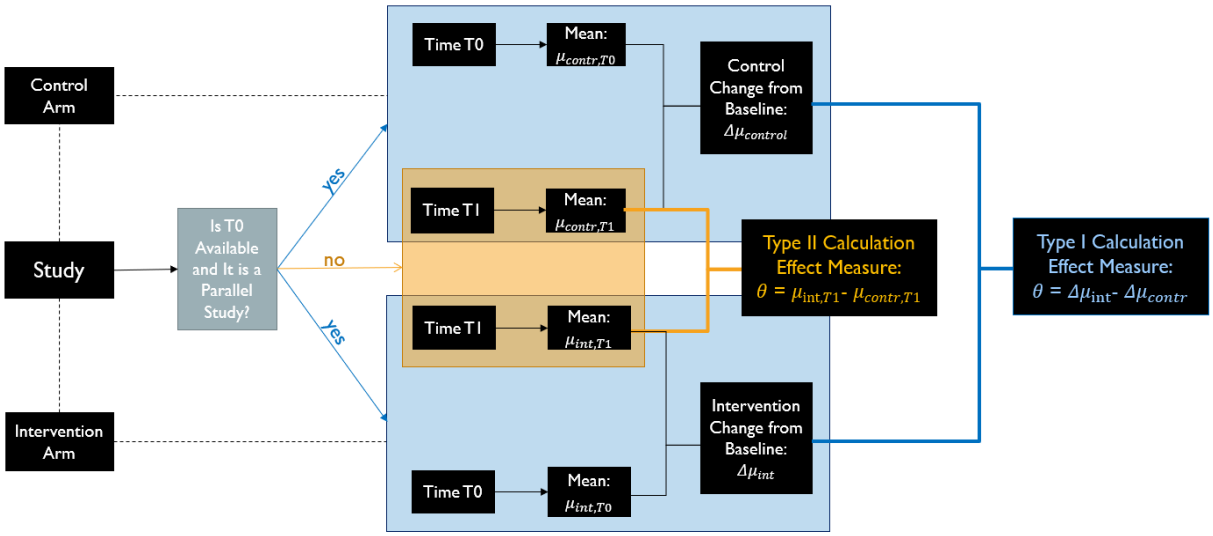


Figure 1: Effect measure calculation scheme

First, the change from baseline and its SE were computed within each arm according to the formula below:

$$\boldsymbol{\Delta\mu_j = \mu_{j,T1} - \mu_{j,T0}} \qquad \textbf{Equation 12}$$

$$\boldsymbol{n_{\Delta\mu_j} = (n_{j,T1} - n_{j,T0})/2} \qquad \textbf{Equation 13}$$

$$\boldsymbol{SE_{\Delta\mu_j} = \sqrt{(n_{j,T0})^2 + (n_{j,T1})^2 - 2 \cdot c \cdot SE_{j,T0} \cdot SE_{j,T1}}} \qquad \textbf{Equation 14}$$

where $\Delta\mu_j$ is the change from baseline mean value, $n_{\Delta\mu_j}$ is the sample size, $SE_{\Delta\mu_j}$ is the standard error, $j \in \{control, intervent\}$ and c is the correlation coefficient that expresses the association of repeated measurements on the same individual at different times.

The mean effect value (θ), and the respective sample size (n_θ) and standard error (SE_θ), were calculated for each study. For endpoints related to ectopic fat for which it was not possible to harmonise all the measurement units, a standardise mean difference procedure (Higgins et al., 2020) was implemented, whereas for the other endpoints the mean effect calculation as described by Crippa and Orsini was applied (Crippa and Orsini, 2016).

For the mean difference calculation two cases were considered depending on whether:

1) the endpoint measurement was available both at baseline, T0, and at the end of the intervention, T1 (parallel studies);
2) the endpoint measurement was available at T1 only (parallel and cross-over studies).

In case (1) the mean effect was obtained as the difference of the change from baseline of the two arms. In case (2) the mean effect was computed as the mean difference between arms at T1. The formulas for the mean effect calculation are provided below:

$$\theta = \begin{cases} \Delta\mu_{int} - \Delta\mu_{contr} & if\ \ \mu_{j,T0} \neq 0\ and\ parallel\ design\ (:= 1) \\ \mu_{int,T1} - \mu_{contr,T1} & otherwise\ (:= 2) \end{cases} \qquad \text{Equation 15}$$

$$n_\theta = \begin{cases} \frac{n_{\Delta\mu_{int}} + n_{\Delta\mu_{contr}}}{2} & if\ (1) \\ \frac{n_{int,T1} + n_{contr,T1}}{2} & if\ (2) \end{cases}$$

Equation 16

$$\boldsymbol{SE_\theta} = \begin{cases} \sqrt{\left[\frac{1}{n_{\Delta\mu_{contr}}} + \frac{1}{n_{\Delta\mu_{int}}}\right] \cdot \frac{SD_{\Delta\mu_{contr}}(n_{\Delta\mu_{contr}} - 1) + SD_{\Delta\mu_{int}}(n_{\Delta\mu_{int}} - 1)}{\sqrt{n_{\Delta\mu_{contr}} + n_{\Delta\mu_{int}} - 2}}} & if(1) \\ \sqrt{\left[\frac{1}{n_{contr,T1}} + \frac{1}{n_{int,T1}} - 2\frac{c_c}{\sqrt{n_{contr,T1} n_{int,T1}}}\right] \cdot \frac{SD_{contr,T1}(n_{contr,T1} - 1) + SD_{int,T1}(n_{int,T1} - 1)}{\sqrt{n_{contr,T1} + n_{int,T1} - 2}}} & if(2) \end{cases}$$

Equation 17

where c_c is the correlation coefficient for cross-over design studies and $SD_{j,T1}$is the standard deviation (SD) of the j-th arm at time T1 for each $j \in \{control, intervent\}$.

For the endpoints for which the unit measure standardisation was not possible, the Hedges standardised mean difference ($\boldsymbol{\theta_S}$) was calculated (Higgins et al., 2020). The approach used was to normalise the mean effect dividing by a pooled SD adjusted with the Hedges Correction factor (HFC):

$$\boldsymbol{\theta_S = HFC \cdot \frac{\theta}{SD_{pooled}}}$$

Equation 18

where $\boldsymbol{\theta}$ is the mean effect calculated following Equation 15 and

$$HFC = \begin{cases} 1 - \frac{3}{4(n_{\Delta\mu_{contr}} + n_{\Delta\mu_{int}} - 2) - 1} & if(1) \\ 1 - \frac{3}{4(n_{contr,T1} + n_{int,T1} - 2) - 1} & if(2) \end{cases}$$

Equation 19

$$SD_{pooled} = \begin{cases} \sqrt{\frac{SD_{\Delta\mu_{contr}}(n_{\Delta\mu_{contr}} - 1) + SD_{\Delta\mu_{int}}(n_{\Delta\mu_{int}} - 1)}{n_{\Delta\mu_{contr}} + n_{\Delta\mu_{int}} - 2}} & if(1) \\ \sqrt{\frac{SD_{contr,T1}(n_{contr,T1} - 1) + SD_{int,T1}(n_{int,T1} - 1)}{n_{contr,T1} + n_{int,T1} - 2}} & if(2) \end{cases}$$

Equation 20

1.1.3. Data imputation

Since in some studies the VP or sample size were not reported, they were imputed.

Three types of imputation were carried out as illustrated in Figure 2.

- Within arm, i.e. if in the same arm of one study a value is missing for T0, but it is present for T1, the T1 value was used as donor, and vice versa.
- Across arms, i.e. within each time value, control values are imputed using intervention as donor or vice versa.
- Across studies, within same arm (control or intervention) and time of measurements (T0 or T1).

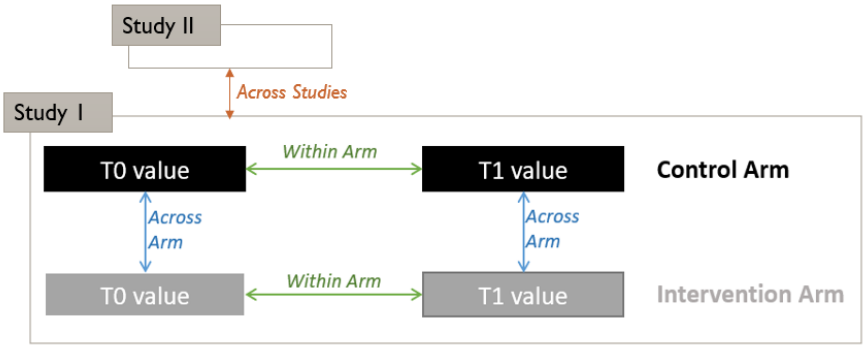


Figure 2: Imputation scheme

First missing sample sizes, then missing standard deviations were imputed. Missing SEs were calculated by applying:

$$SE = \frac{SD}{\sqrt{n}} \qquad \text{Equation 6}$$

to the imputed standard deviations and sample sizes.

When the SD was not available for any of the arms and time points in a study, the method described by Marinho et al. (2003) was used for the imputation. Within the same arm and within the same time point, the log SD was regressed over the log mean value, following the formula:

$$log(SD_{j,Ti}) \sim b0_{j,Ti} + b1_{j,Ti} * log(\mu_{j,Ti}) \qquad \textbf{Equation 21}$$

where $j \in \{control, intervent\}$, $i \in \{0,1\}$. The regression coefficients were estimated using only the studies for which $SD_{j,Ti}$ and $\mu_{j,Ti}$ were not missing.

1.1.4. Correlation coefficient

As anticipated in Equation 14 and Equation 17, correlation coefficients were used to calculate the mean effect. There are two types of correlation:

1) Between baseline (T0) and end-of-treatment (T1) measurements within arms (c).
2) Between intervention and control measurements at T1 for cross-over studies, since the two treatments are administered to the same individuals (c_c).

Given the uncertainty in the level of correlation and the limited evidence that is available to provide an accurate and precise estimate for our BoE, an expert knowledge elicitation (EFSA, 2014) was carried out with the members of the Working Group on sugars to estimate these values.

The following approach was used:

- Identify a range that would cover with 95% probability the true value of the correlation coefficient for all the metabolic endpoints considered after intake of different doses of sugars in a population similar to the target population.
- Identify within the range the 'best estimate' for the correlation coefficient in cross-over studies.

The range and the best estimate were first elicited at individual level and then a consensus was reached by a collegial discussion.

The full procedure including the expert knowledge elicitation (EKE) questions and the related evidence is reported in Appendix C –.

It was decided to use the same correlation coefficient values for parallel and cross-over studies (c_c and c) under the assumption that values elicited considering cross-over trials would represent a lower boundary for parallel studies (i.e. endpoint levels observed at the beginning and at the end of a trial in the same individuals would be similar to endpoint levels observed on the same individuals to which different treatments are administered at different timepoints). The best estimate of the correlation coefficient at the end of the process was $c_c = c = 0.82$.

1.2. Descriptive statistics

The numbers of studies computed for each endpoint are reported in **Table 1** for Q1 and **Table 2** for Q2. Overall, for Q1 there were more studies (232) than for Q2 (82).

Table 1: Table 1. N of studies for Q1 for each endpoint

Endpoint	N of studies
BMI	6
Body Fat	6
Body Weight	11
Clamp (Hepatic)	3
Clamp (Whole Body)	4
Diastolic Blood Pressure	11
Ectopic Fat (Liver Fat)	5
Ectopic Fat (VAT)	4
Fasting Glucose	18
Fasting Insulin	16
HDL Cholesterol	24
LDL Cholesterol	21
OGTT Glucose	11
OGTT Insulin	11
Systolic Blood Pressure	11
Total Cholesterol	29
Fasting triglycerides	29
Uric Acid	8
Waist Circumference	4

Table 2: Table 2. N of studies for Q2 for each endpoint

Endpoint	N of studies
BMI	1
Body Fat	1
Body Weight	2
Clamp (Hepatic)	1
Clamp (Whole Body)	3
Diastolic Blood Pressure	5
Ectopic Fat (Liver Fat)	4
Ectopic Fat (VAT)	2
Fasting Glucose	9
Fasting Insulin	8
HDL Cholesterol	7
LDL Cholesterol	7
OGTT Glucose	3
OGTT Insulin	3
Systolic Blood Pressure	5
Total Cholesterol	7
Fasting triglycerides	9
Uric Acid	5
Waist Circumference	1

1

1.3. Graphical display

Three types of graphical displays were used to describe the data: 1. Forest plots; 2. Scatterplots; and 3. Funnel plots.

Forest plots were used to display the mean effect, SE, sample size and respective 95% CI for each study, along with main study characteristics (i.e. sex, sugars dose as E%, sugars type, subject characteristics, sugars source, type of diet, body weight change, intervention duration in weeks, Risk of Bias and design). The pooled mean effect and its 95% CI are reported for correlation coefficients of 0.82, 0.5 and 0.99. The 95% PI is also provided for the correlation coefficient of 0.82. Finally, the heterogeneity is reported using the I^2 (Higgins and Thompson, 2002). Forest plots can be found in Appendix G of the scientific opinion.

Scatterplots were used to check the relationship between the sugars dose (E%) and the endpoint mean effect and to identify possible explanatory variables for the dose–response model for the endpoints fasting triglycerides (both Q1 and Q2), fasting glucose, fasting insulin and body weight (only Q1). For the latter set of outcomes the number of observations was too small to allow any investigation of dose–response.

In the following scatterplots are reported:

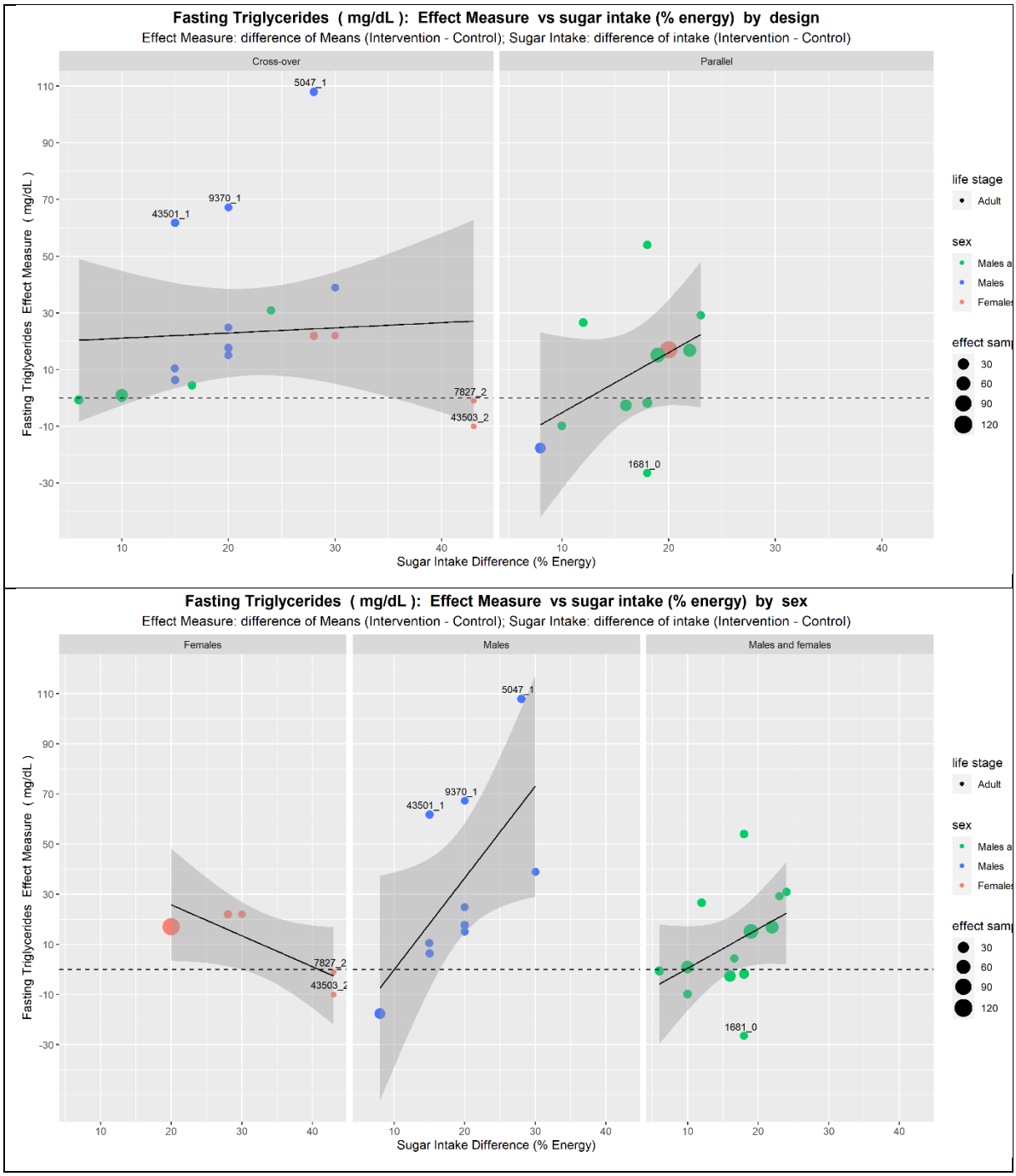
- On the y-axis the endpoint values, respectively fasting triglycerides [mg/dL], fasting glucose [mg/dL], fasting insulin [pmol/L] and body weight [kg].
- On the x-axis the sugars intake difference [E%].
- The variables investigated, respectively:
  - Design: cross-over, parallel
  - Sex: males, females, mixed
  - Diet: *ad libitum*, isocaloric with neutral energy balance (isocaloric neutral), isocaloric with positive energy balance (isocaloric positive)
  - Source of sugars: foods, beverages, mixed
  - Wash-out linked to design (cross-over with wash-out, cross-over without wash-out, parallel)
  - Run-in
  - Duration (≤ 8 weeks, > 8 weeks)
  - Type of sugar (fructose, glucose, mixed fructose and glucose)
  - Body weight effect change (≤ −2, [−2, 2], >2)
  - Risk of bias – RoB (Tier 1 and Tier 2).
- Mean study values are characterised by sex (different colours), effect sample size $\boldsymbol{n_\theta}$ (size of the points) and life stage (studies in children or adults, identified by the shape of the point).

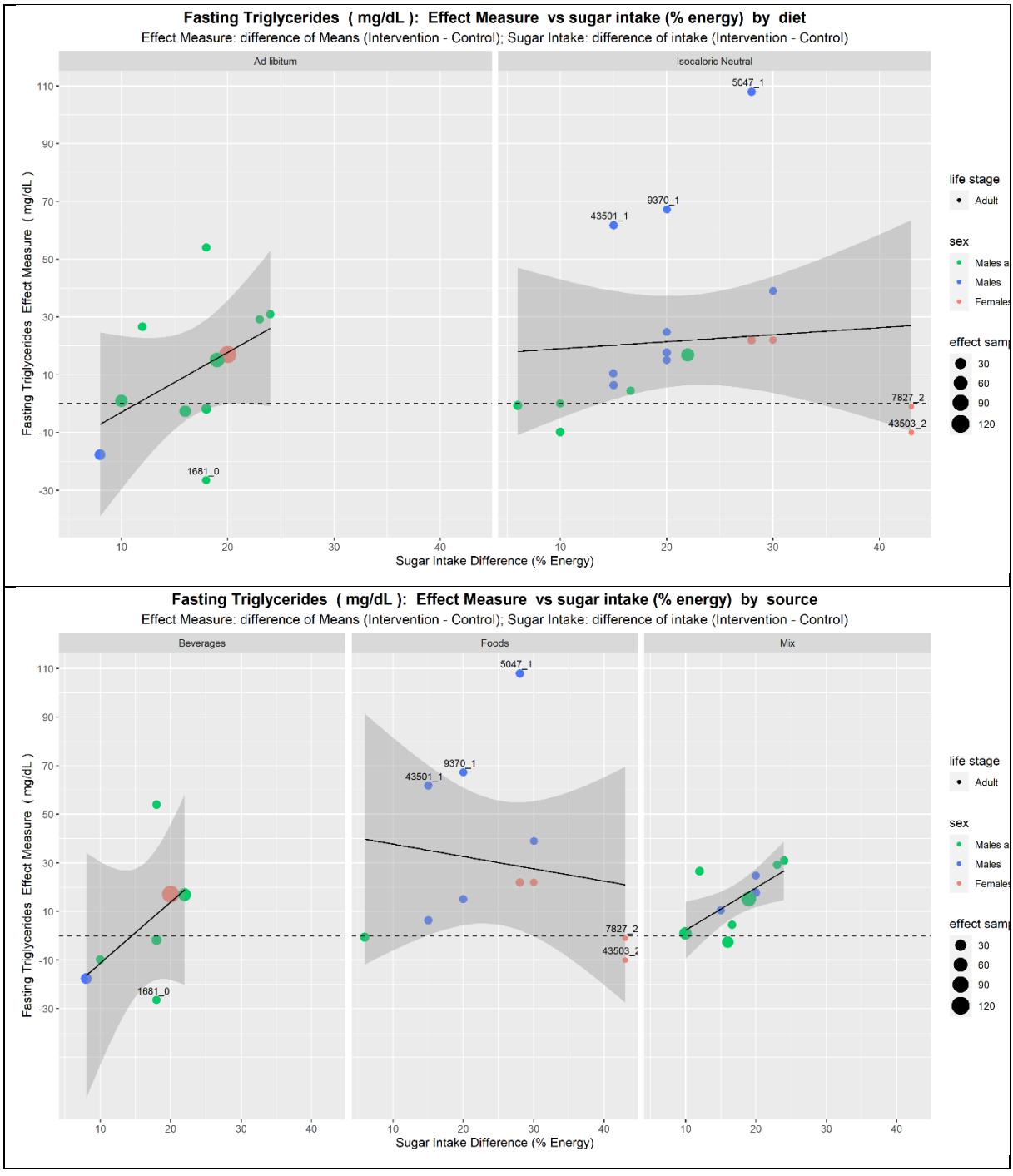
When needed, continuous variables were categorised (i.e. body weight effect change and duration). In these scatterplots an unadjusted linear regression across sugars intake difference ($\boldsymbol{d}$) and mean effect ($\boldsymbol{\theta}$) are displayed by categories of the explanatory variables. Data were not meta-analysed for the visual inspection. The variables indicating a possible modification of the relationship, both regression slope or intercept, were investigated further in the dose–response analysis.

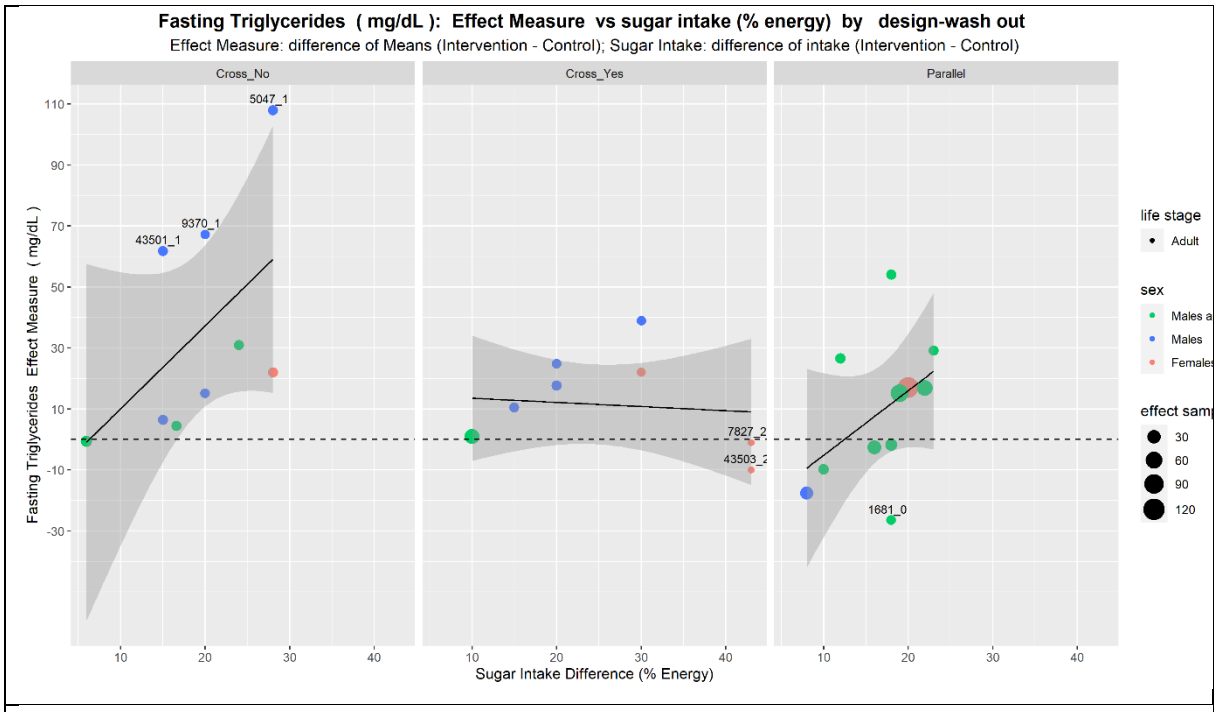
Funnel plots were used to explore whether publication bias could have occurred (Light and Pillemer, 1984). The y-axis represents study precision (i.e. SE of the mean effect) and the x-axis shows the study's mean effect. An asymmetric funnel indicates a relationship between the mean effect estimate and study precision: this suggests the possibility of either publication bias or a systematic difference between studies with higher and lower precision (highly correlated to study size). Studies falling out of the funnel might be indicative of publication bias. To support the assessment of possible publication bias, different colours were used to identify funnels corresponding to various levels of statistical

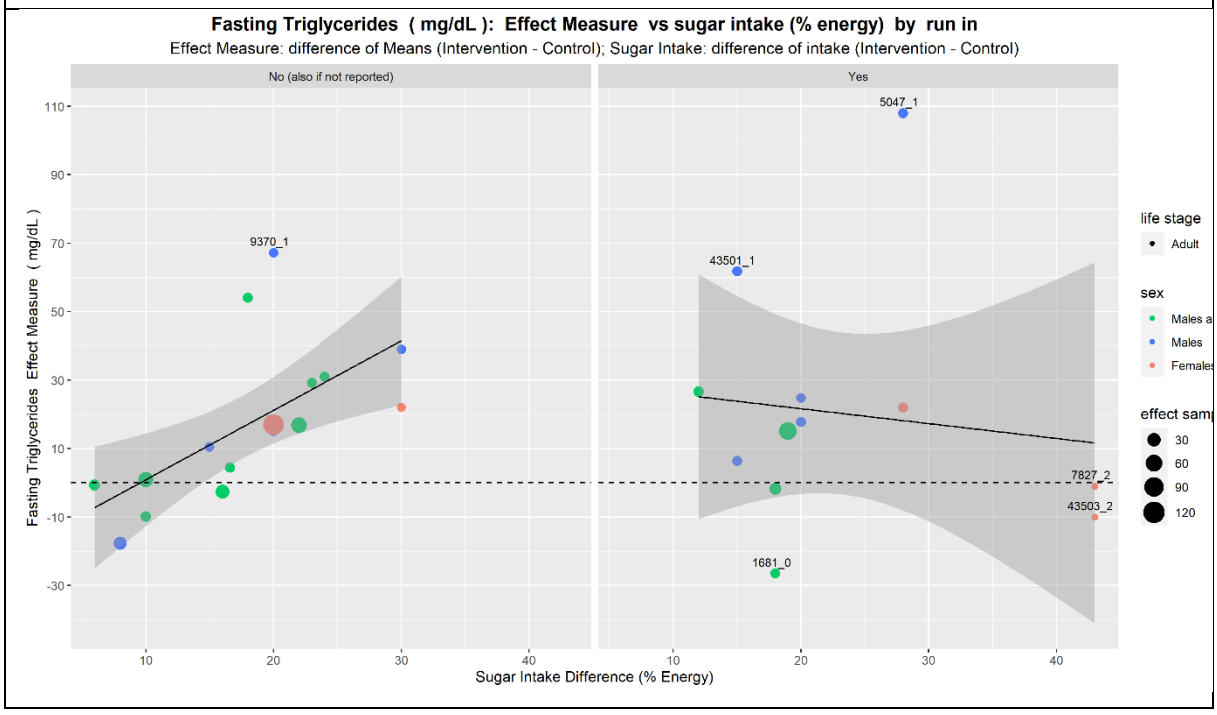
significance of the effects (the lighter the colour the higher the significance level). They can be found in Appendix H of the scientific opinion.

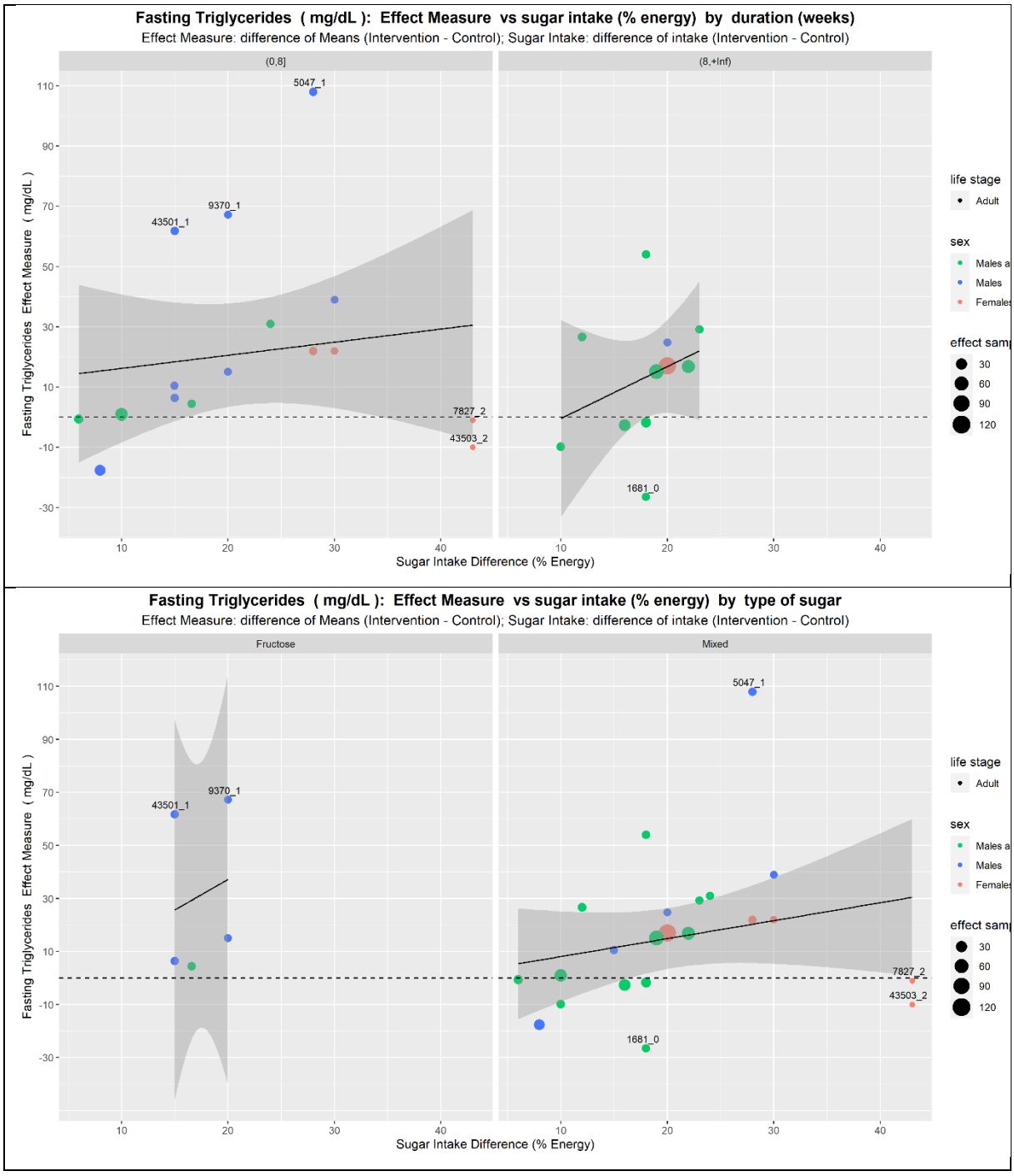
1.3.1. Fasting triglycerides (Q1)




Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by diet
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Ad libitum
Isocaloric Neutral
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
43503_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120
Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by source
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Beverages
Foods
Mix
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by design-wash out
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross_No
Cross_Yes
Parallel
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
43503_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by run in
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
No (also if not reported)
Yes
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
43503_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by duration (weeks)
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
(0,8]
(8,+Inf)
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
43503_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120
Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by type of sugar
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Fructose
Mixed
Fasting Triglycerides Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
5047_1
9370_1
43501_1
7827_2
43503_2
1681_0
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120

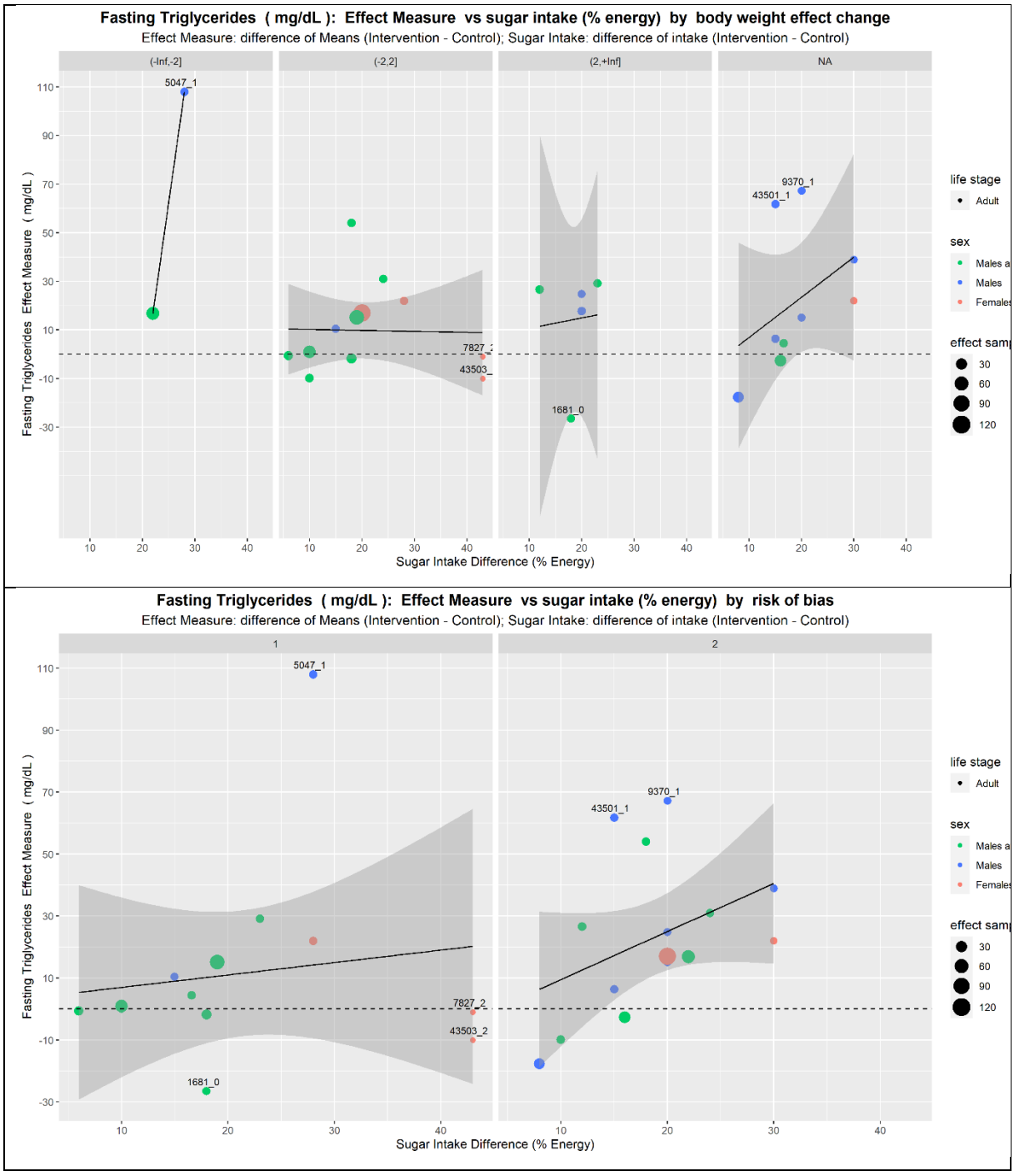
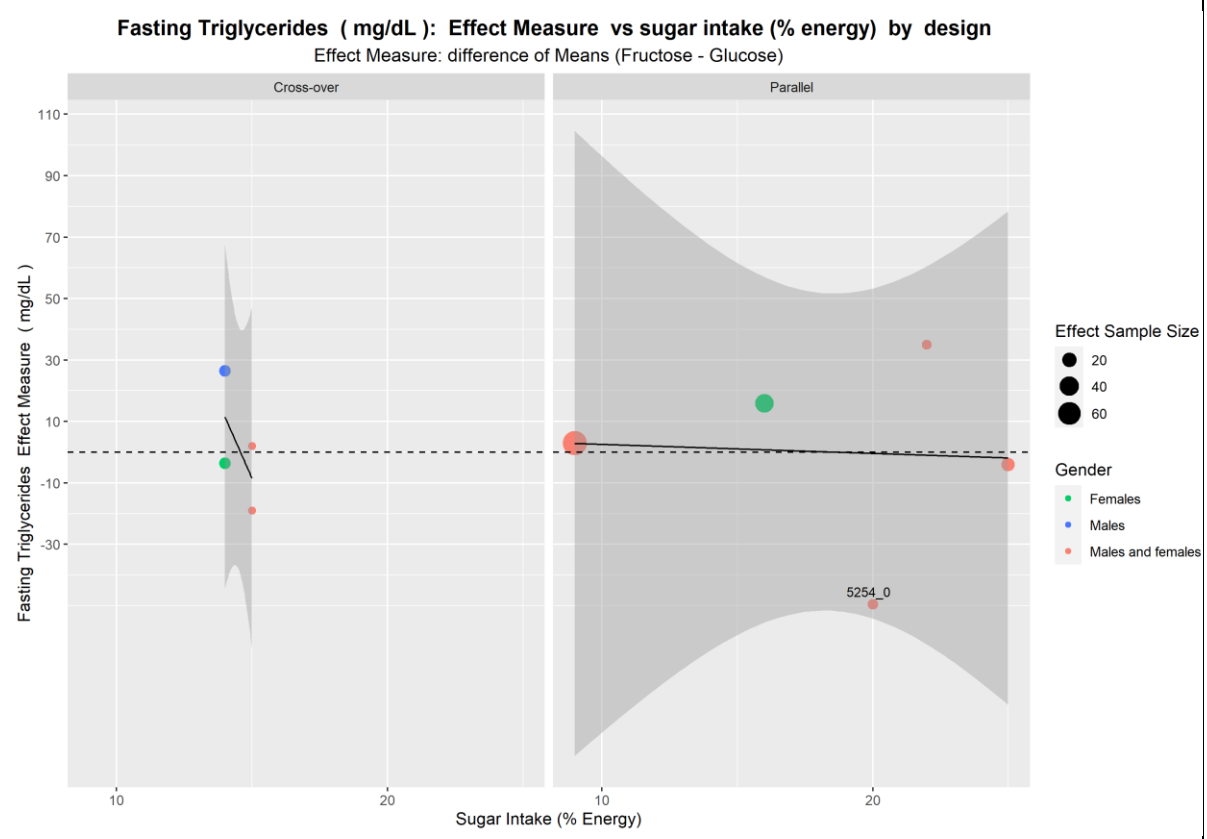
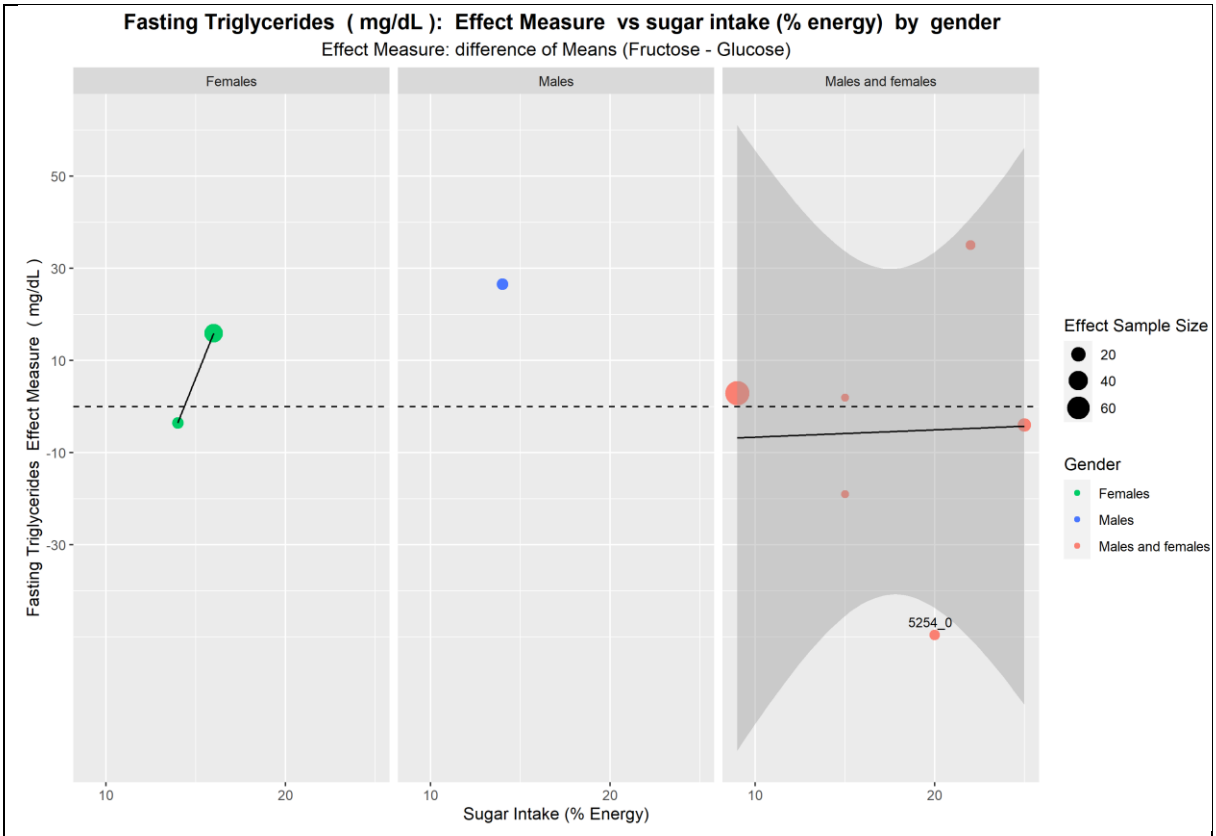
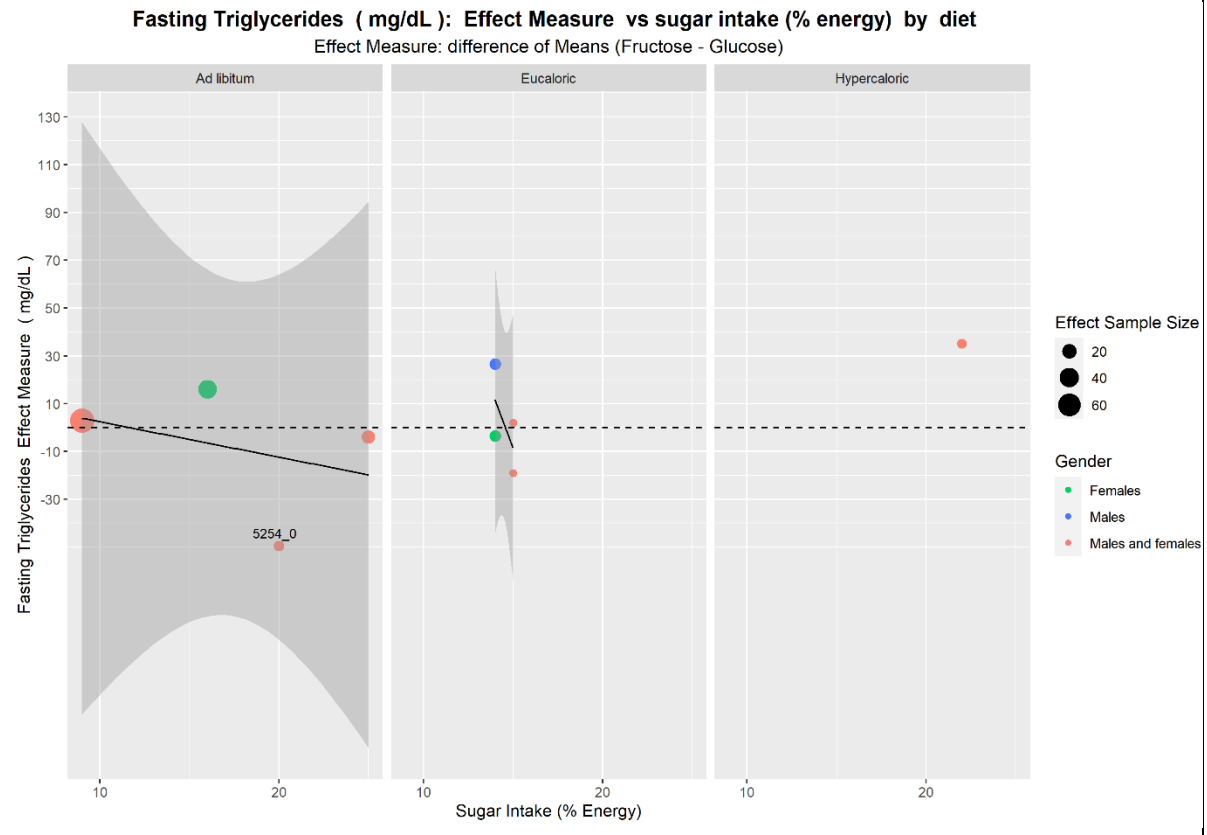


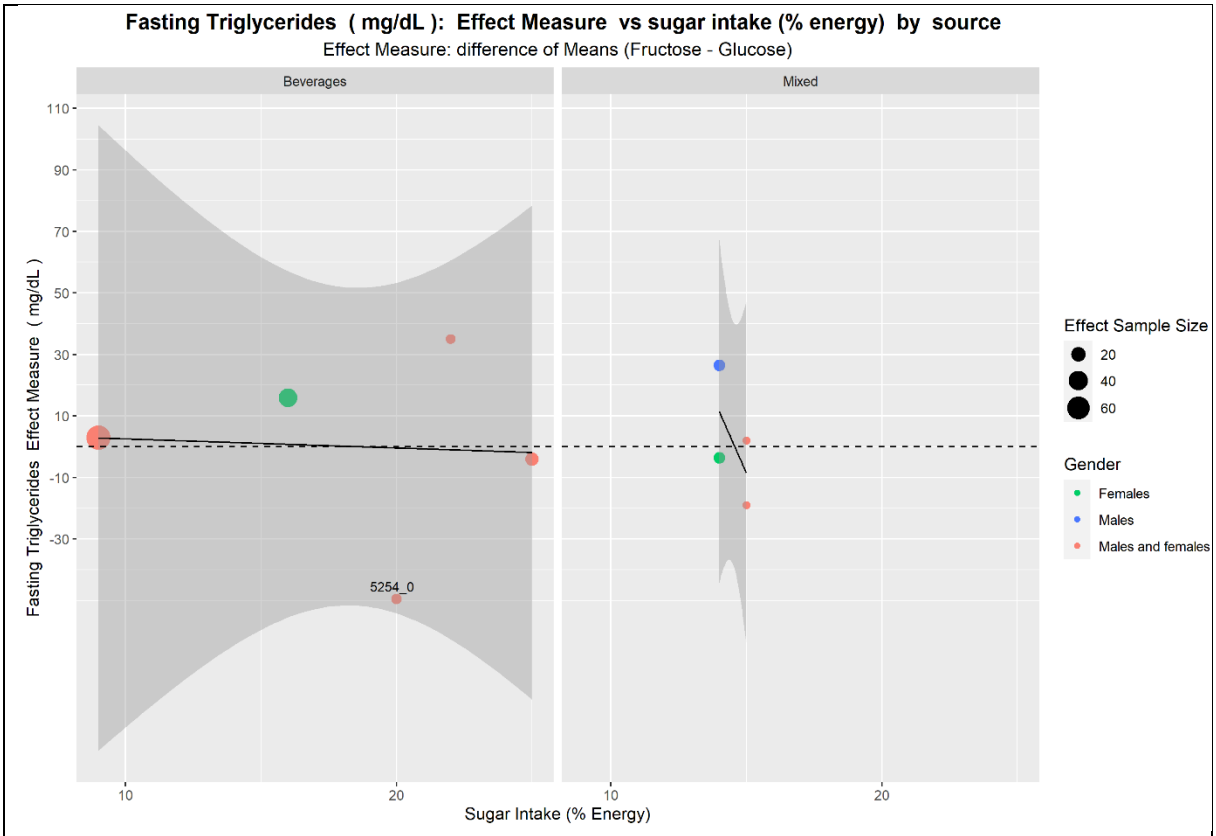
Figure 3: Scatterplots and linear regression between fasting triglycerides mean effect (Q1) and sugar intake arm difference (data not meta-analysed)

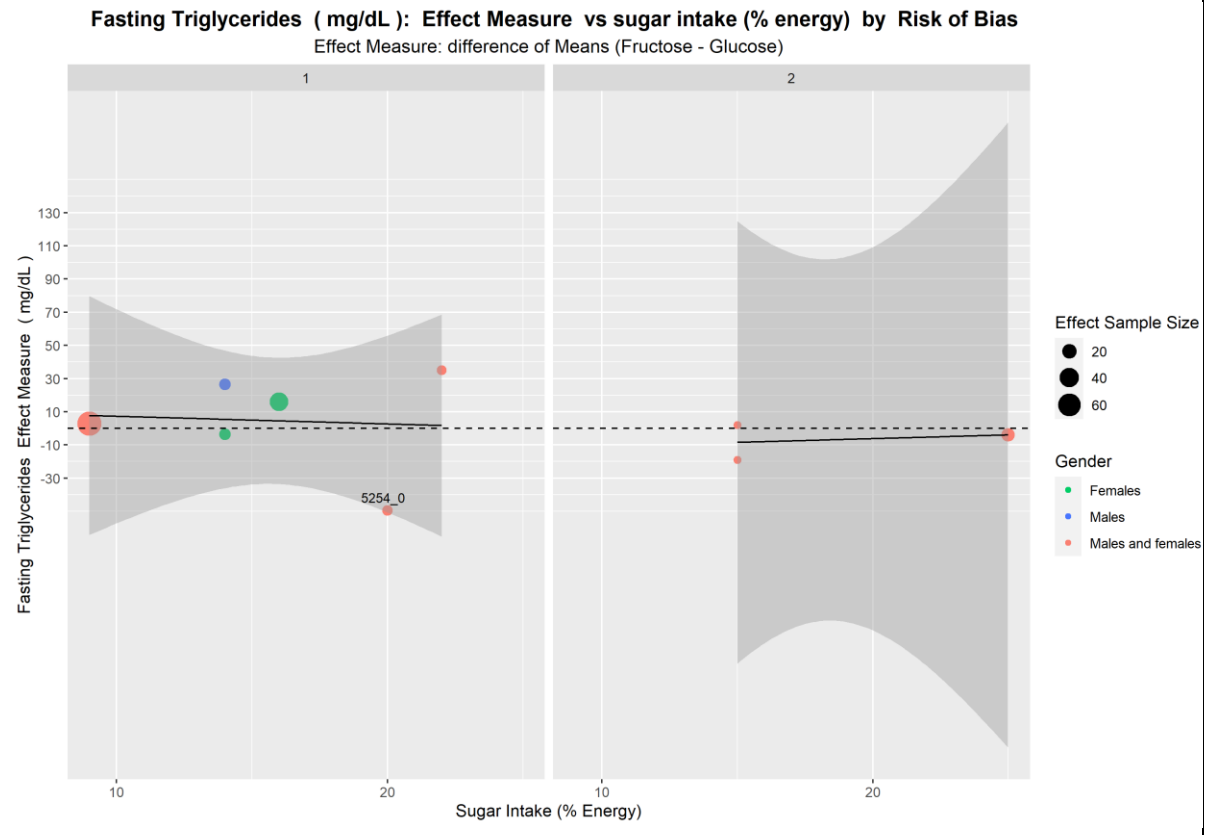
1.3.2. Fasting triglycerides (Q2)

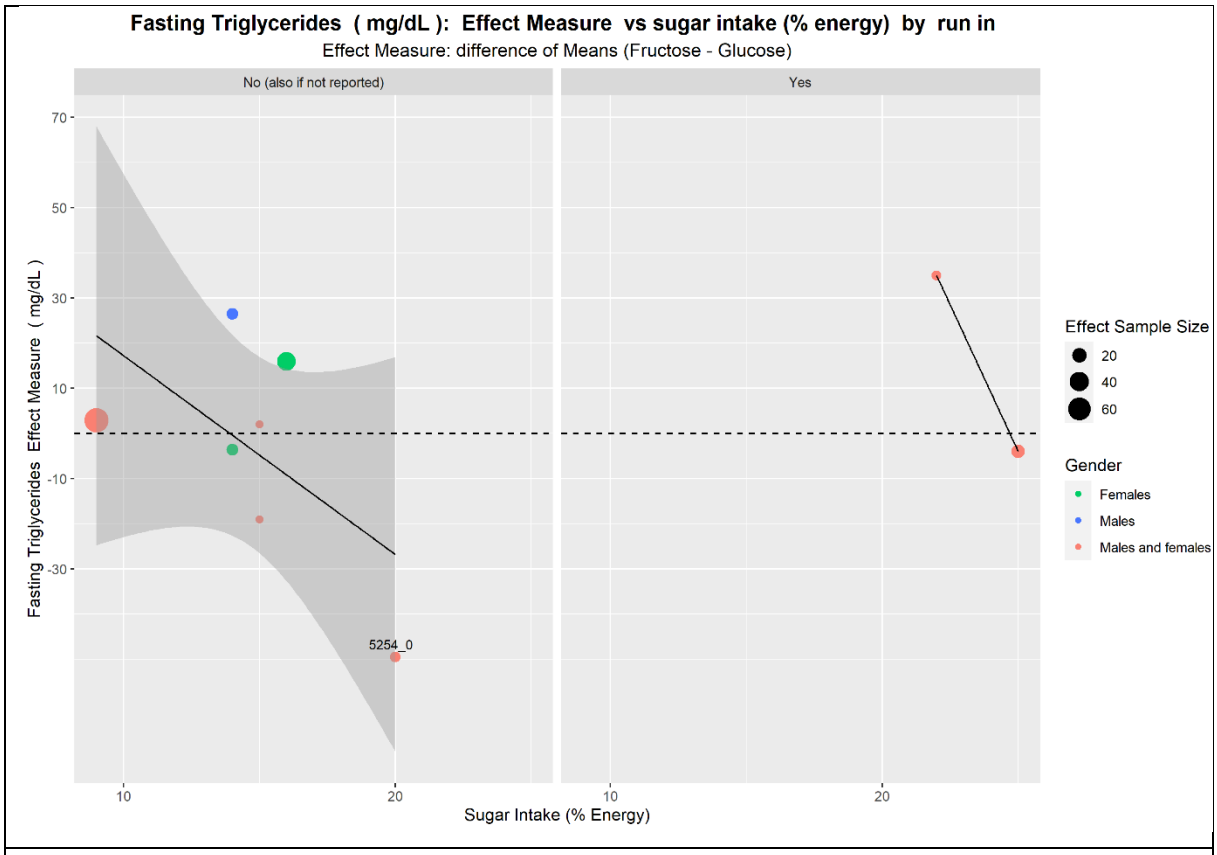


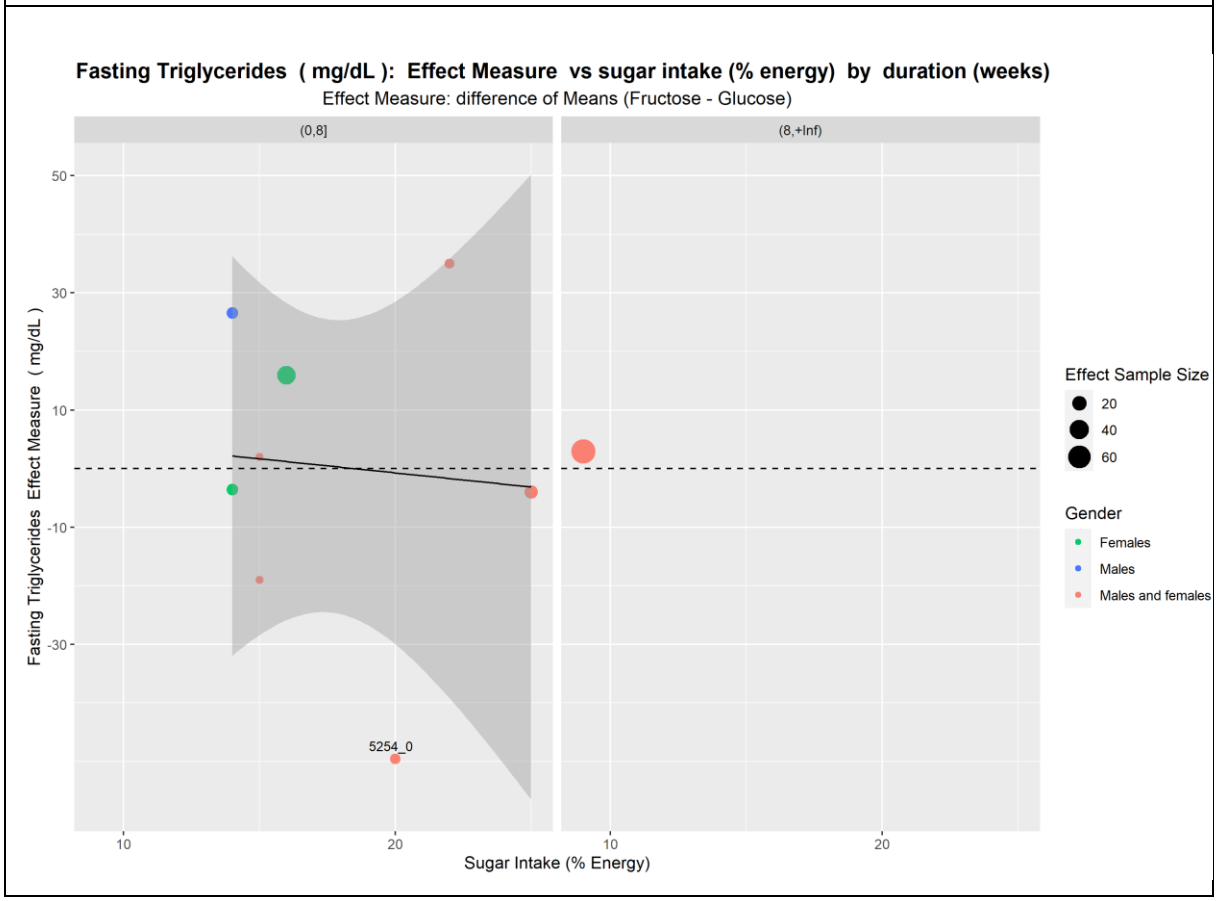

Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by gender
Effect Measure: difference of Means (Fructose - Glucose)
Females
Males
Males and females
Fasting Triglycerides Effect Measure (mg/dL)
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by diet
Effect Measure: difference of Means (Fructose - Glucose)
Ad libitum
Eucaloric
Hypercaloric
Fasting Triglycerides Effect Measure (mg/dL)
130
110
90
70
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by source
Effect Measure: difference of Means (Fructose - Glucose)
Beverages
Mixed
Fasting Triglycerides Effect Measure (mg/dL)
110
90
70
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by Risk of Bias
Effect Measure: difference of Means (Fructose - Glucose)
1
2
Fasting Triglycerides Effect Measure (mg/dL)
130
110
90
70
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by run in
Effect Measure: difference of Means (Fructose - Glucose)
No (also if not reported)
Yes
Fasting Triglycerides Effect Measure (mg/dL)
70
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females


Fasting Triglycerides (mg/dL): Effect Measure vs sugar intake (% energy) by duration (weeks)
Effect Measure: difference of Means (Fructose - Glucose)
(0,8]
(8,+Inf)
Fasting Triglycerides Effect Measure (mg/dL)
50
30
10
-10
-30
5254_0
10
20
Sugar Intake (% Energy)
Effect Sample Size
20
40
60
Gender
Females
Males
Males and females

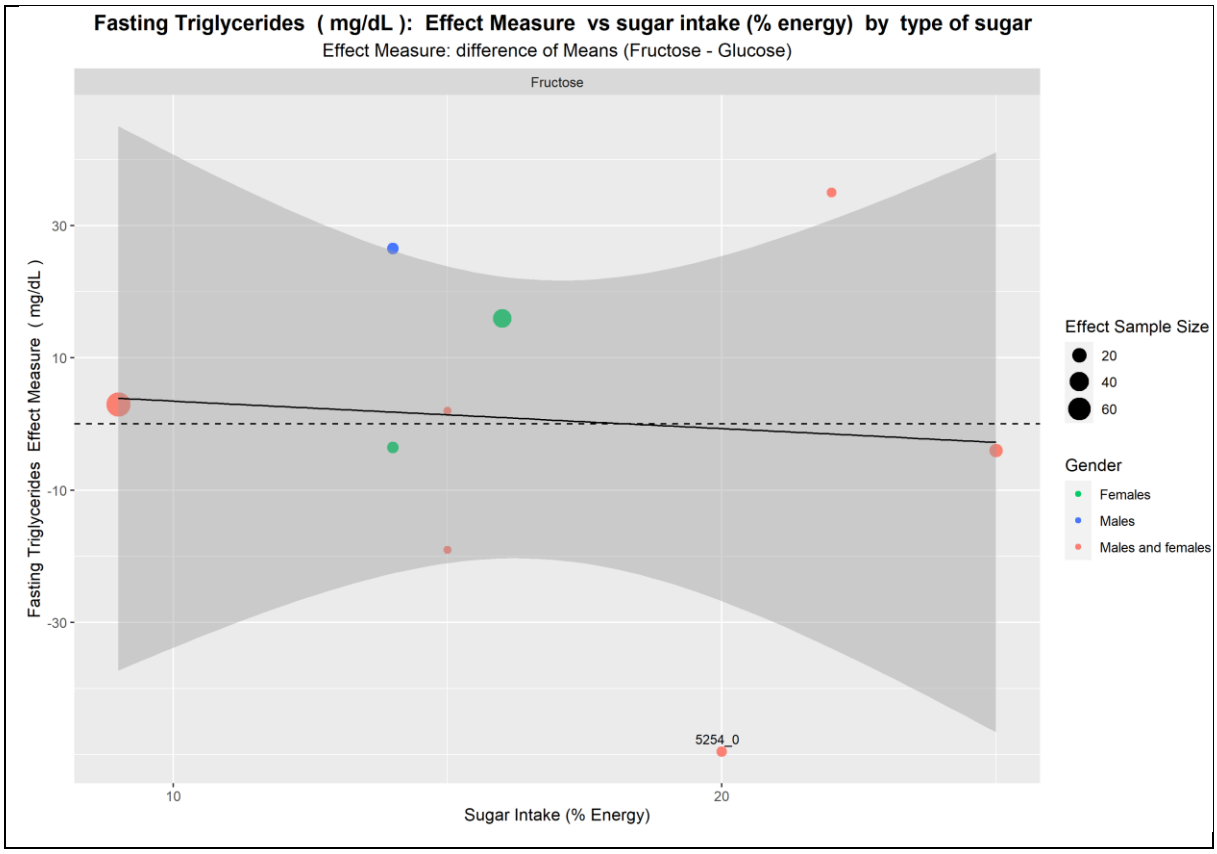
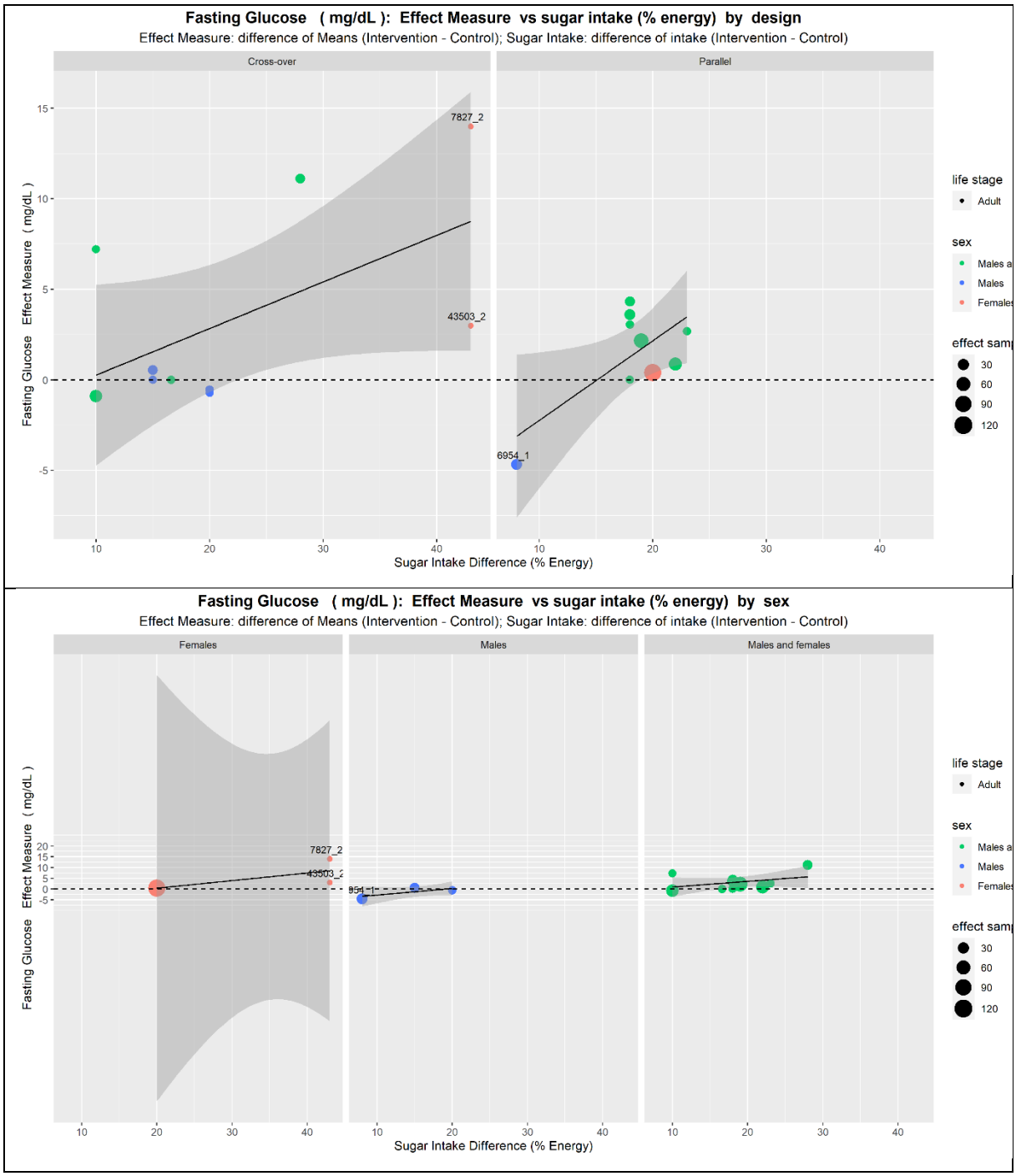
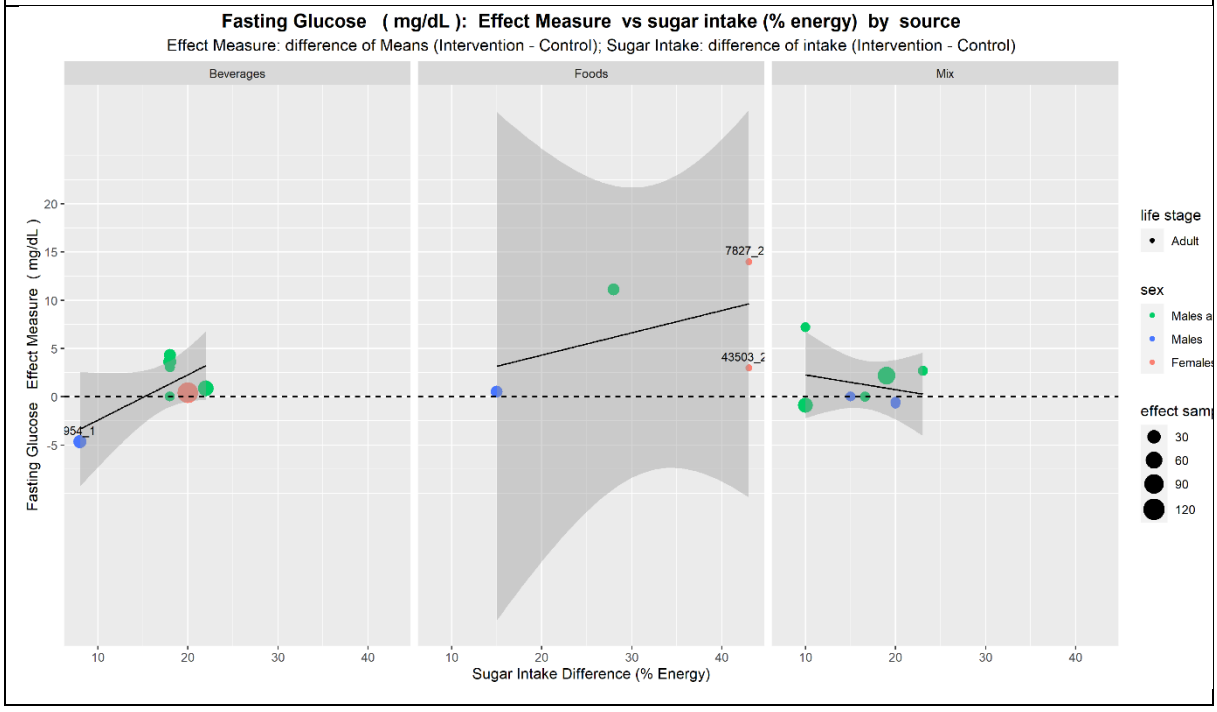


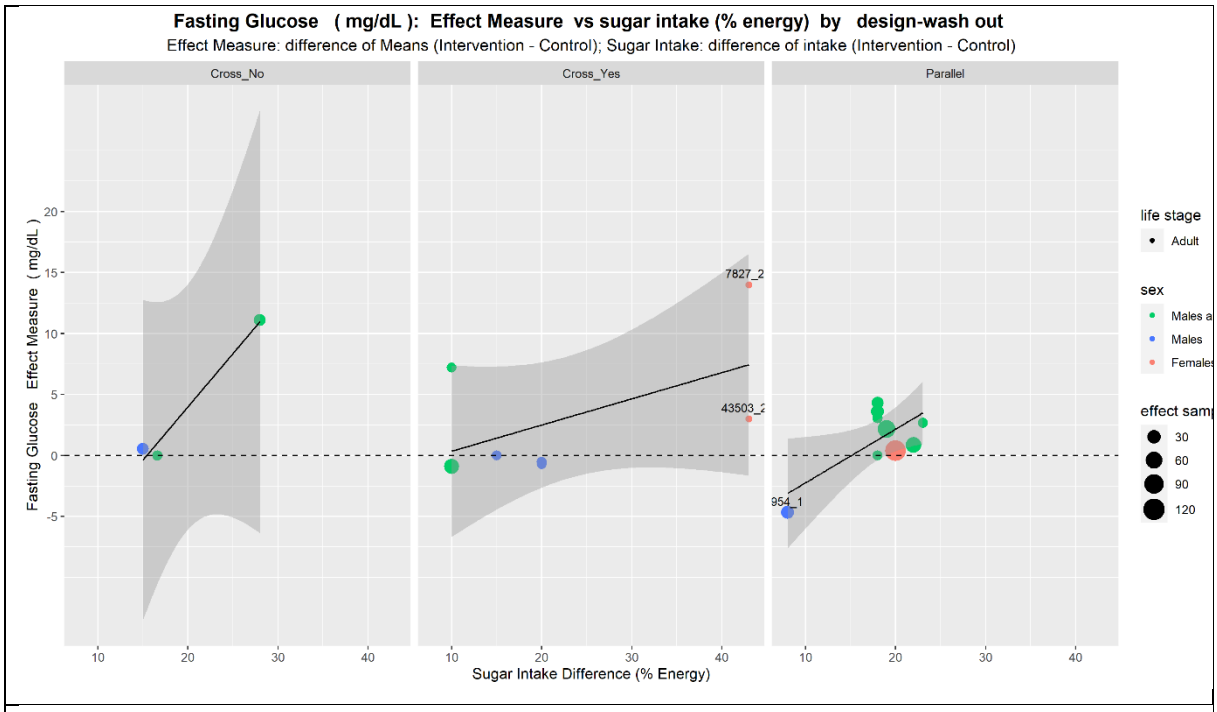
Figure 4: Scatterplots and linear regression between fasting triglycerides mean effect (Q2) and sugar intake arm difference (data not meta-analysed)

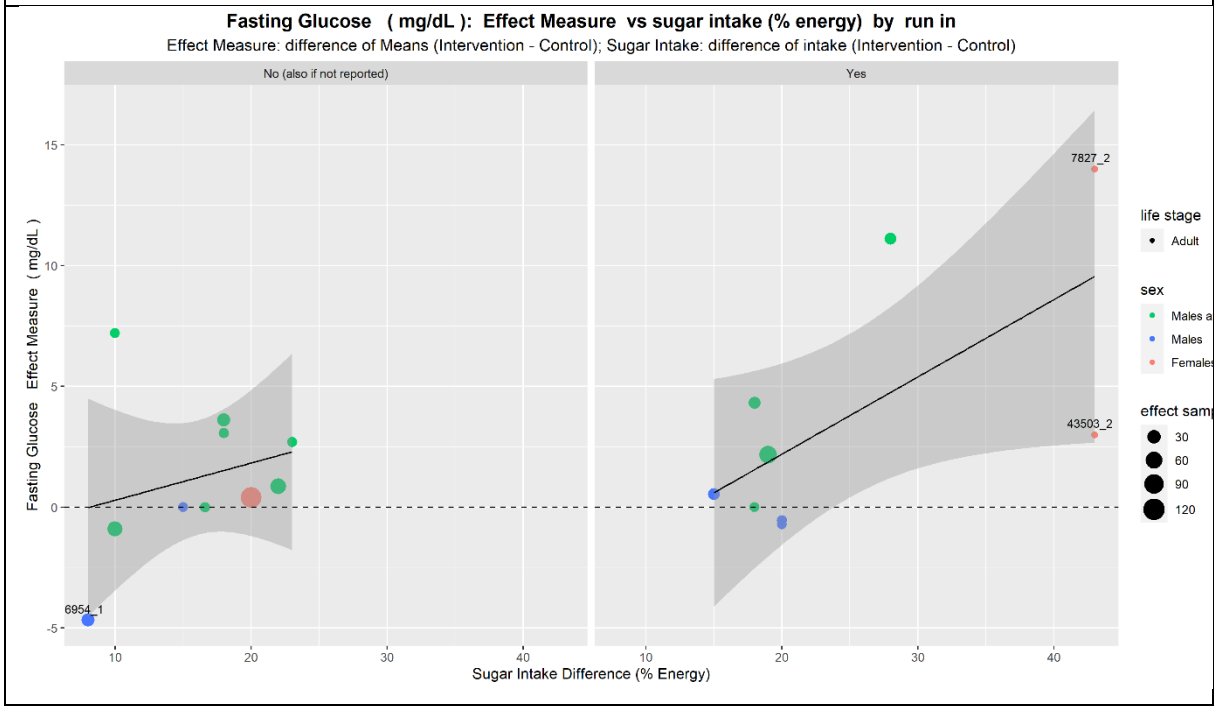
1.3.3. Fasting glucose

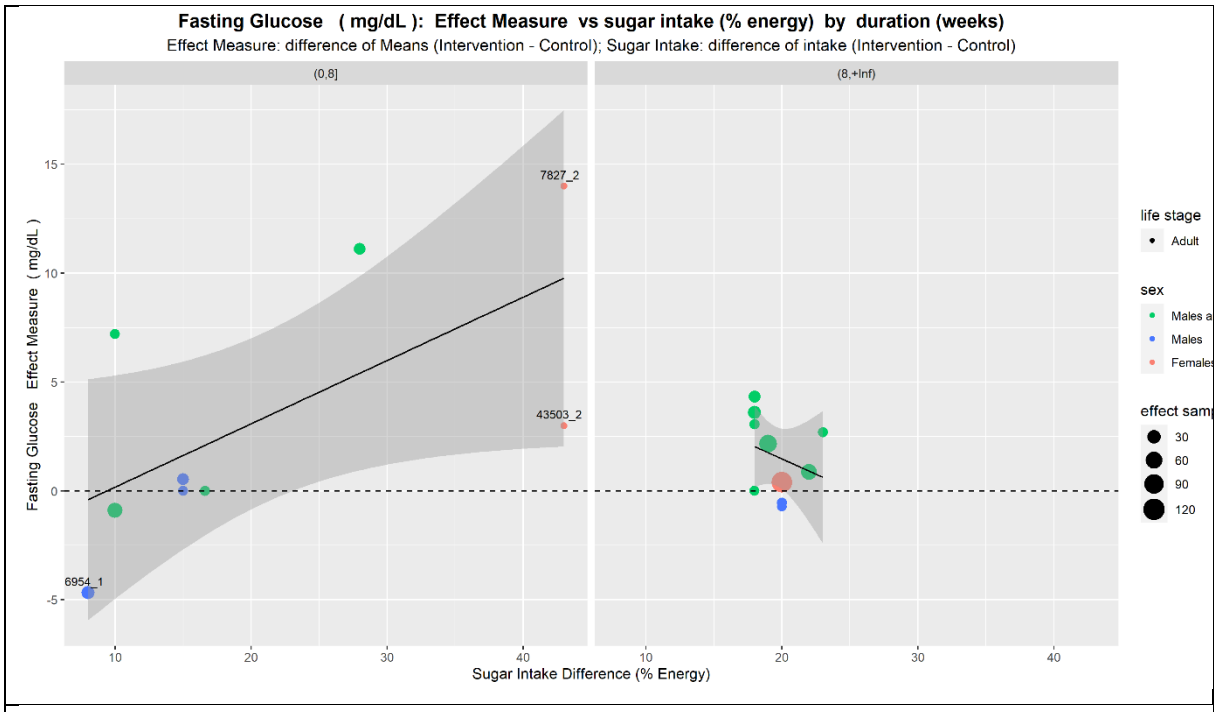


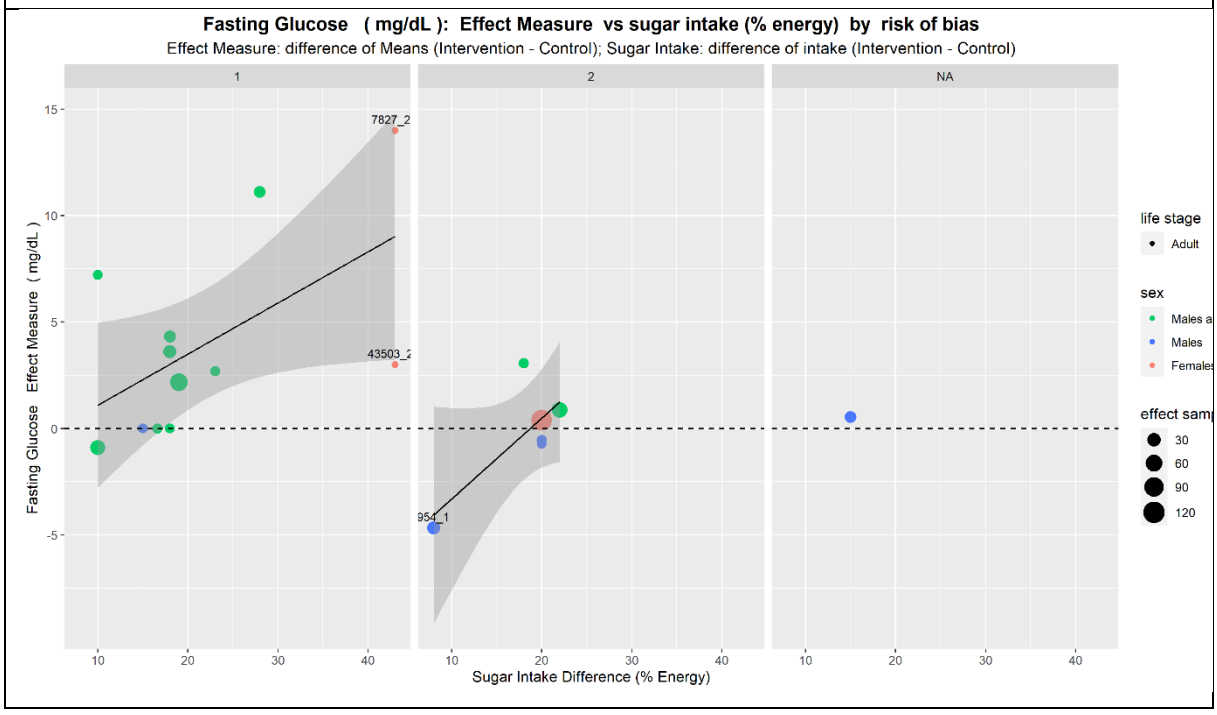

Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by diet
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Ad libitum
Isocaloric Neutral
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
7827_2
43503_2
6954_1
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by source
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Beverages
Foods
Mix
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
7827_2
43503_2
954_1
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by design-wash out
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross_No
Cross_Yes
Parallel
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
7827_2
43503_2
954_1
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by run in
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
No (also if not reported)
Yes
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
7827_2
43503_2
6954_1
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by duration (weeks)
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
(0,8]
(8,+Inf)
7827_2
43503_2
6954_1
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120


Fasting Glucose (mg/dL): Effect Measure vs sugar intake (% energy) by risk of bias
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
1
2
NA
7827_2
43503_2
954_1
Fasting Glucose Effect Measure (mg/dL)
Sugar Intake Difference (% Energy)
life stage
Adult
sex
Males a
Males
Females
effect sam
30
60
90
120

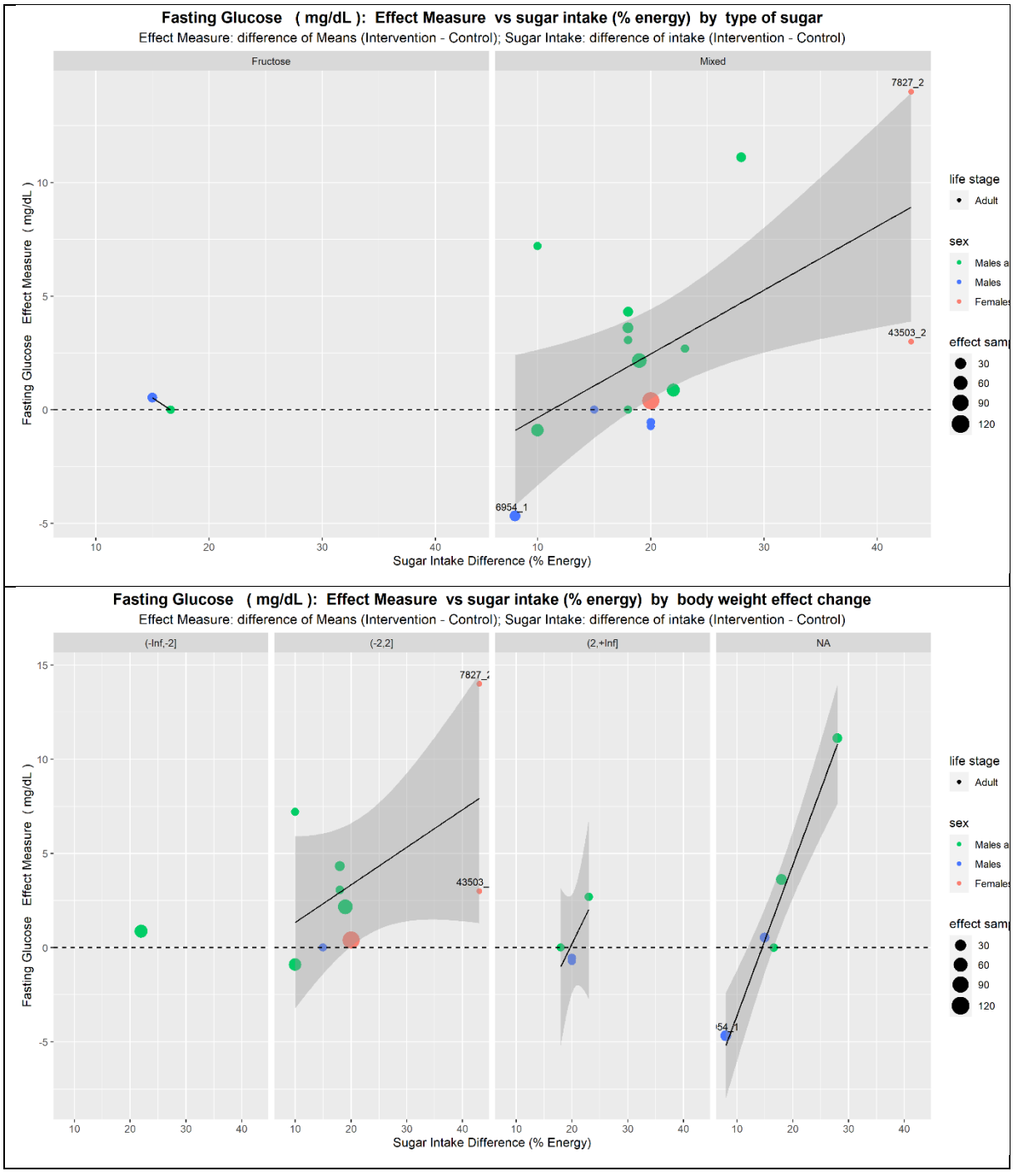
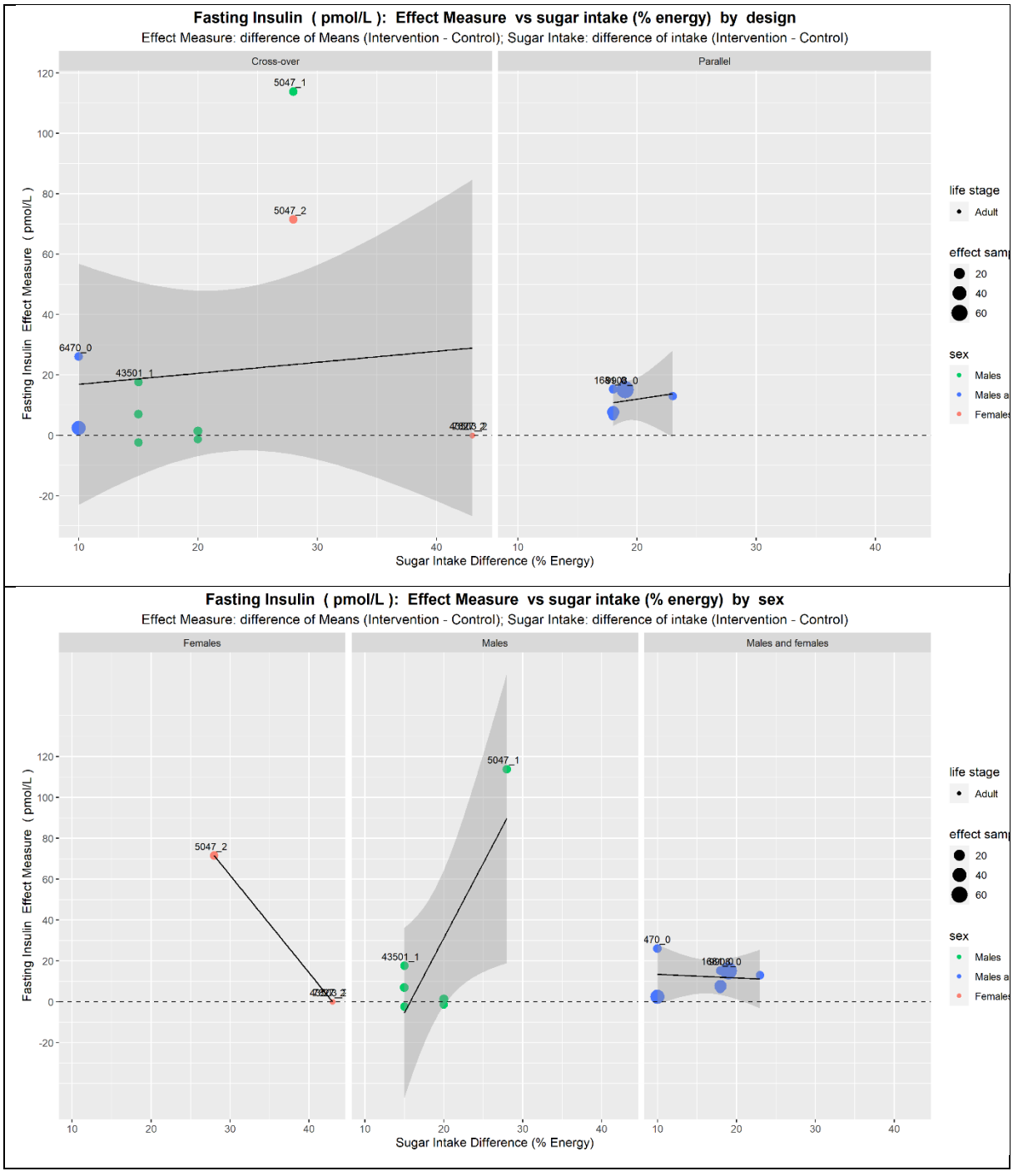
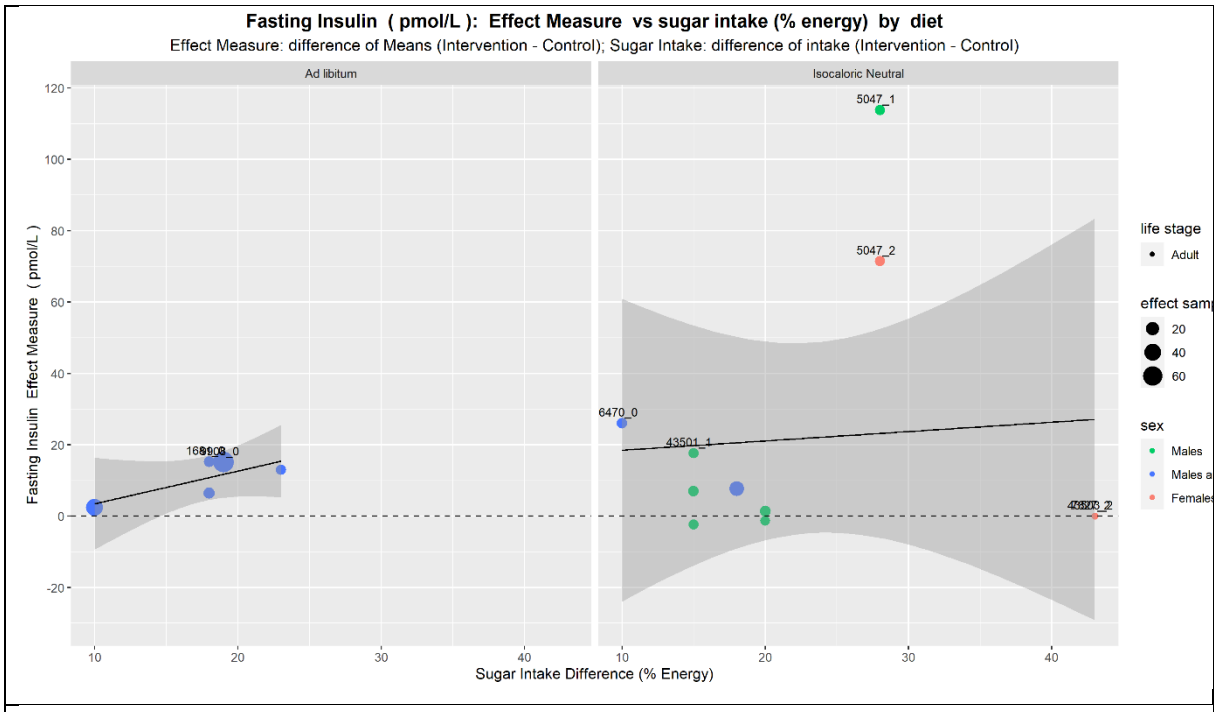
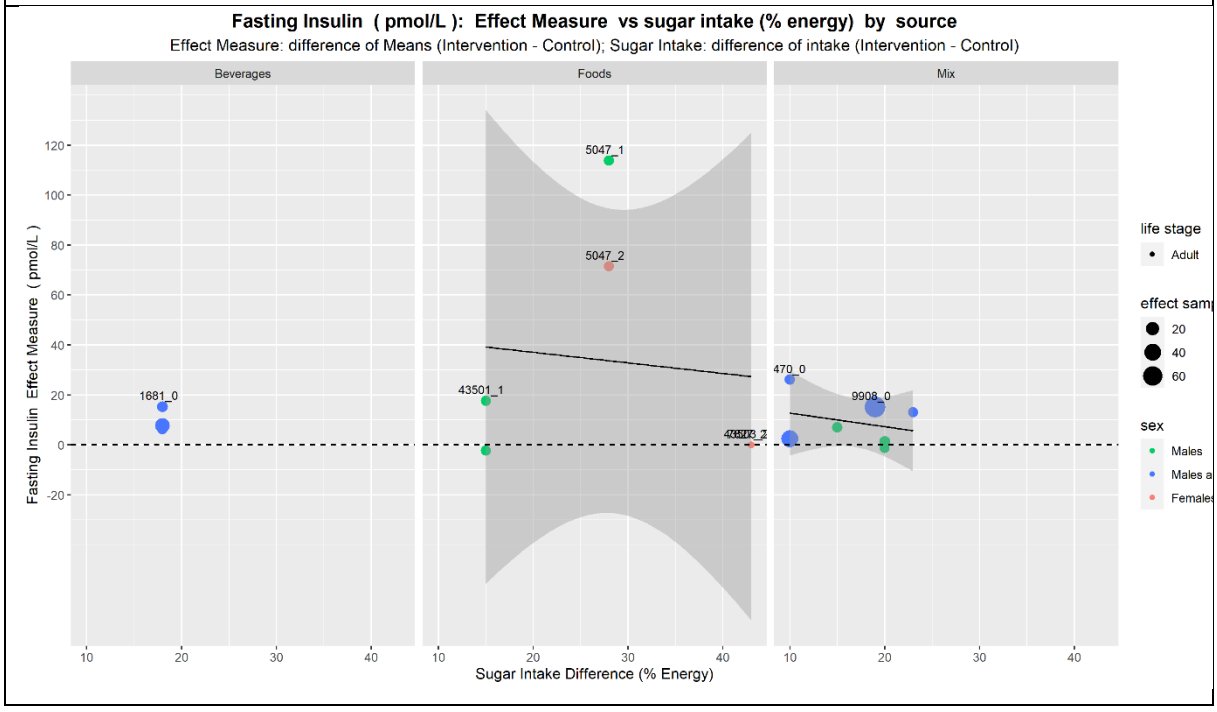


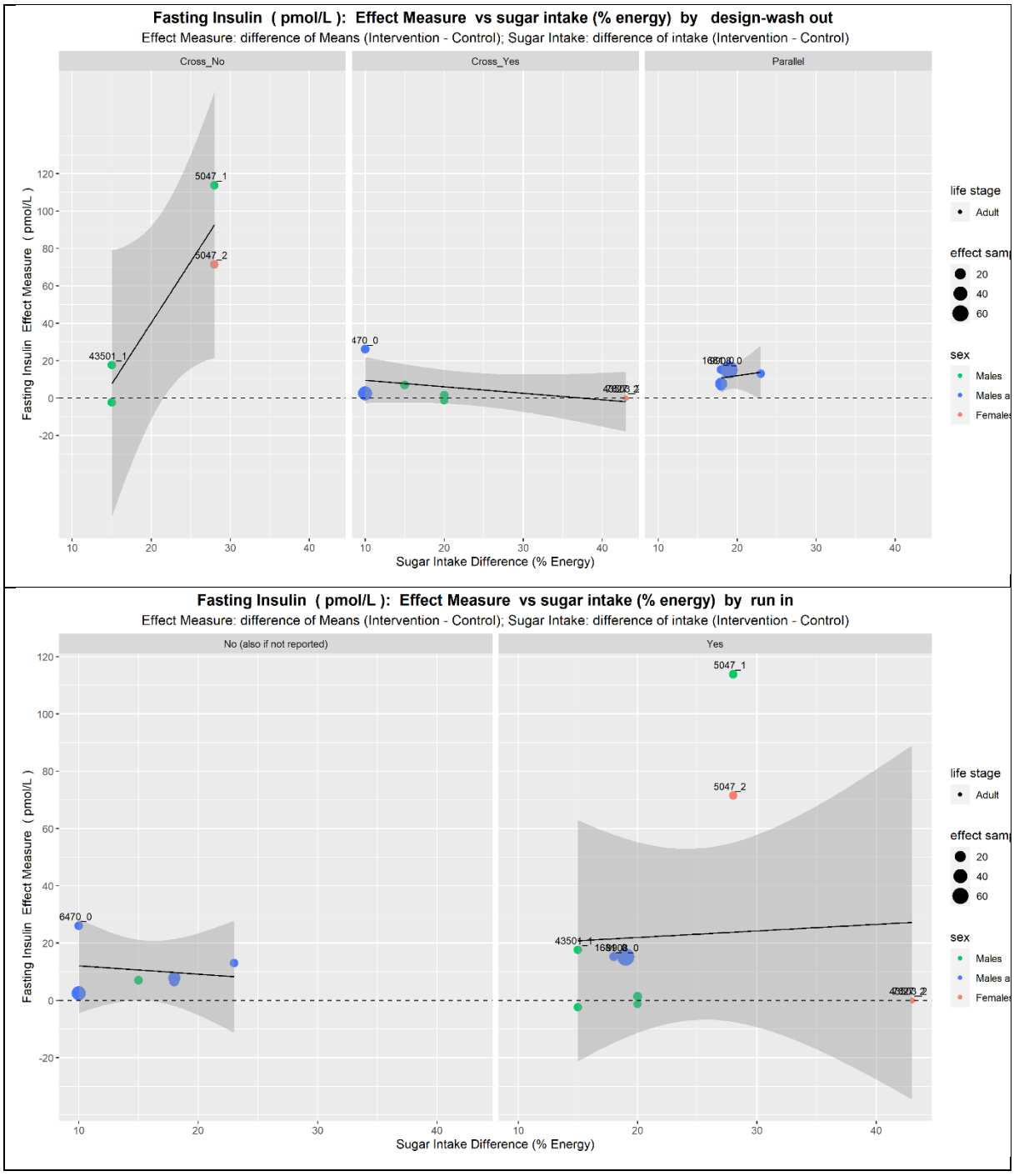
Figure 5: Scatterplots and linear regression between fasting glucose mean effect (Q1) and sugar intake arm difference (data not meta-analysed)

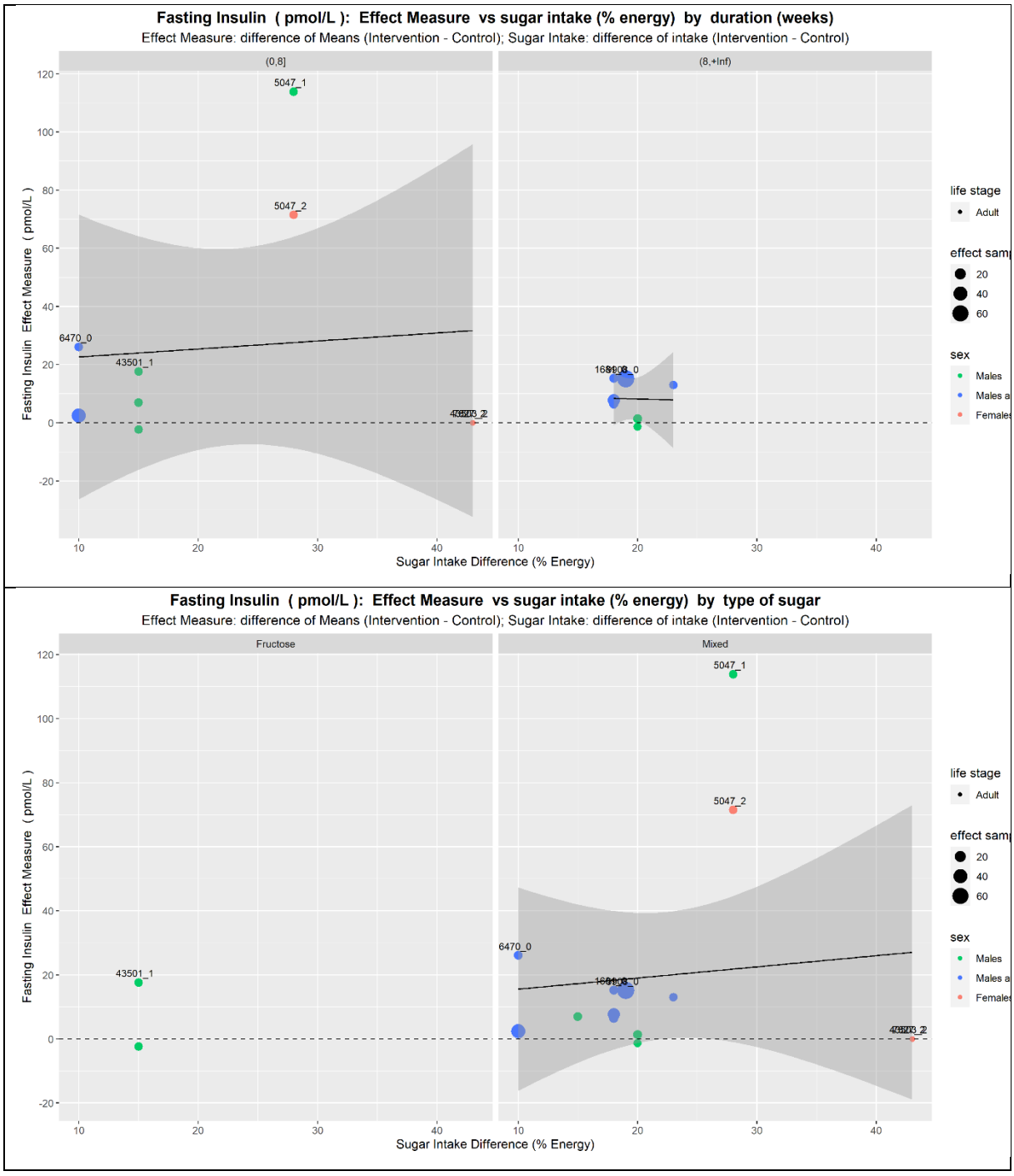
1.3.4. Fasting insulin


Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by design
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross-over
Parallel
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
6470_0
43501_1
life stage
Adult
effect sam
20
40
60
sex
Males
Females
Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by sex
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Females
Males
Males and females
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
43501_1
470_0
life stage
Adult
effect sam
20
40
60
sex
Males
Females


Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by diet
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Ad libitum
Isocaloric Neutral
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
6470_0
43501_1
life stage
Adult
effect sam
20
40
60
sex
Males
Females


Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by source
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Beverages
Foods
Mix
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
1681_0
43501_1
470_0
9908_0
life stage
Adult
effect sam
20
40
60
sex
Males
Females


Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by design-wash out
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross_No
Cross_Yes
Parallel
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
43501_1
470_0
life stage
Adult
effect sam
20
40
60
sex
Males
Males a
Females
Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by run in
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
No (also if not reported)
Yes
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
6470_0
43501_1
life stage
Adult
effect sam
20
40
60
sex
Males
Males a
Females


Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by duration (weeks)
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
(0,8]
(8,+Inf)
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
6470_0
43501_1
life stage
Adult
effect sam
20
40
60
sex
Males
Males a
Females
Fasting Insulin (pmol/L): Effect Measure vs sugar intake (% energy) by type of sugar
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Fructose
Mixed
Fasting Insulin Effect Measure (pmol/L)
Sugar Intake Difference (% Energy)
5047_1
5047_2
6470_0
43501_1
life stage
Adult
effect sam
20
40
60
sex
Males
Males a
Females

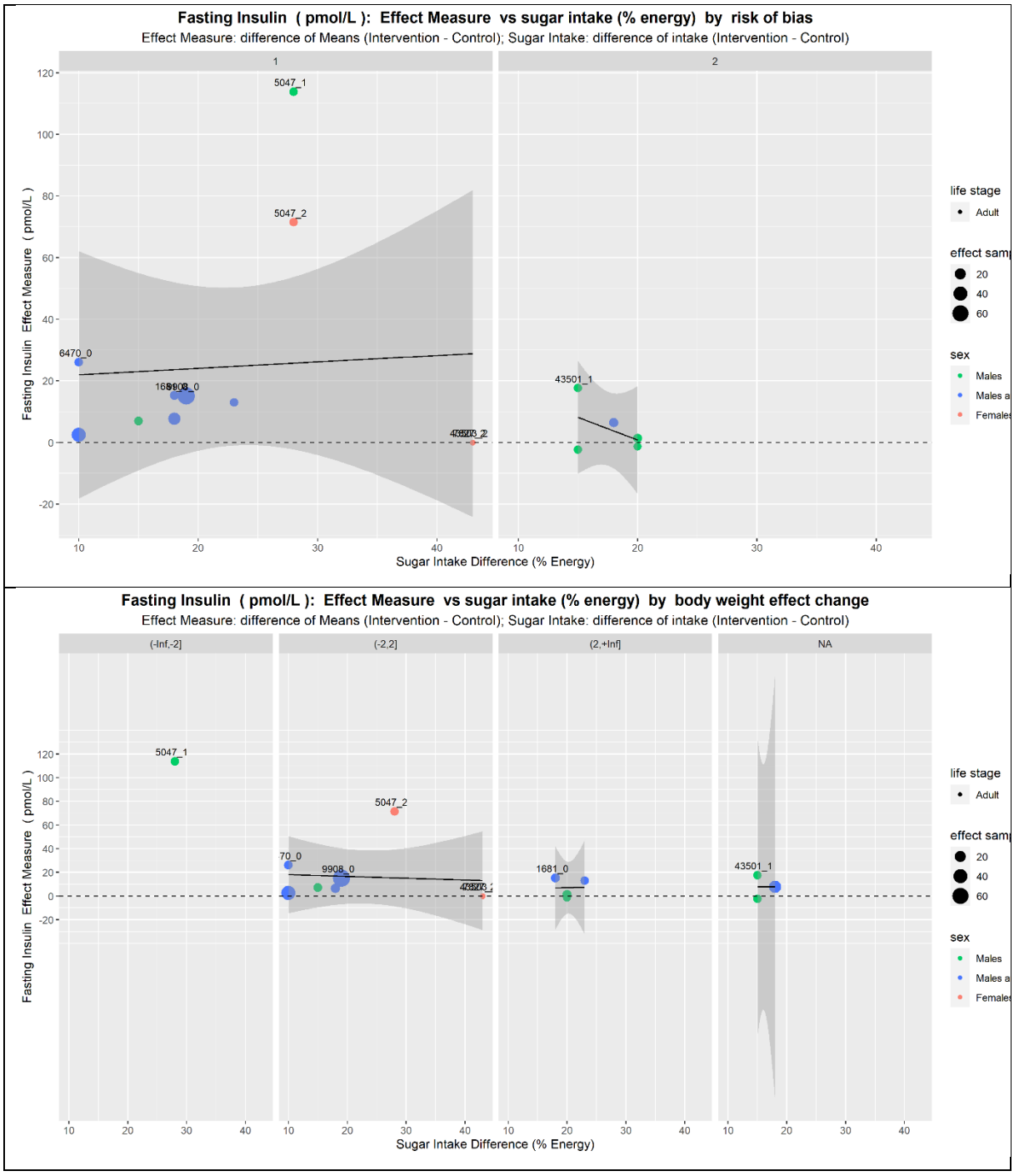
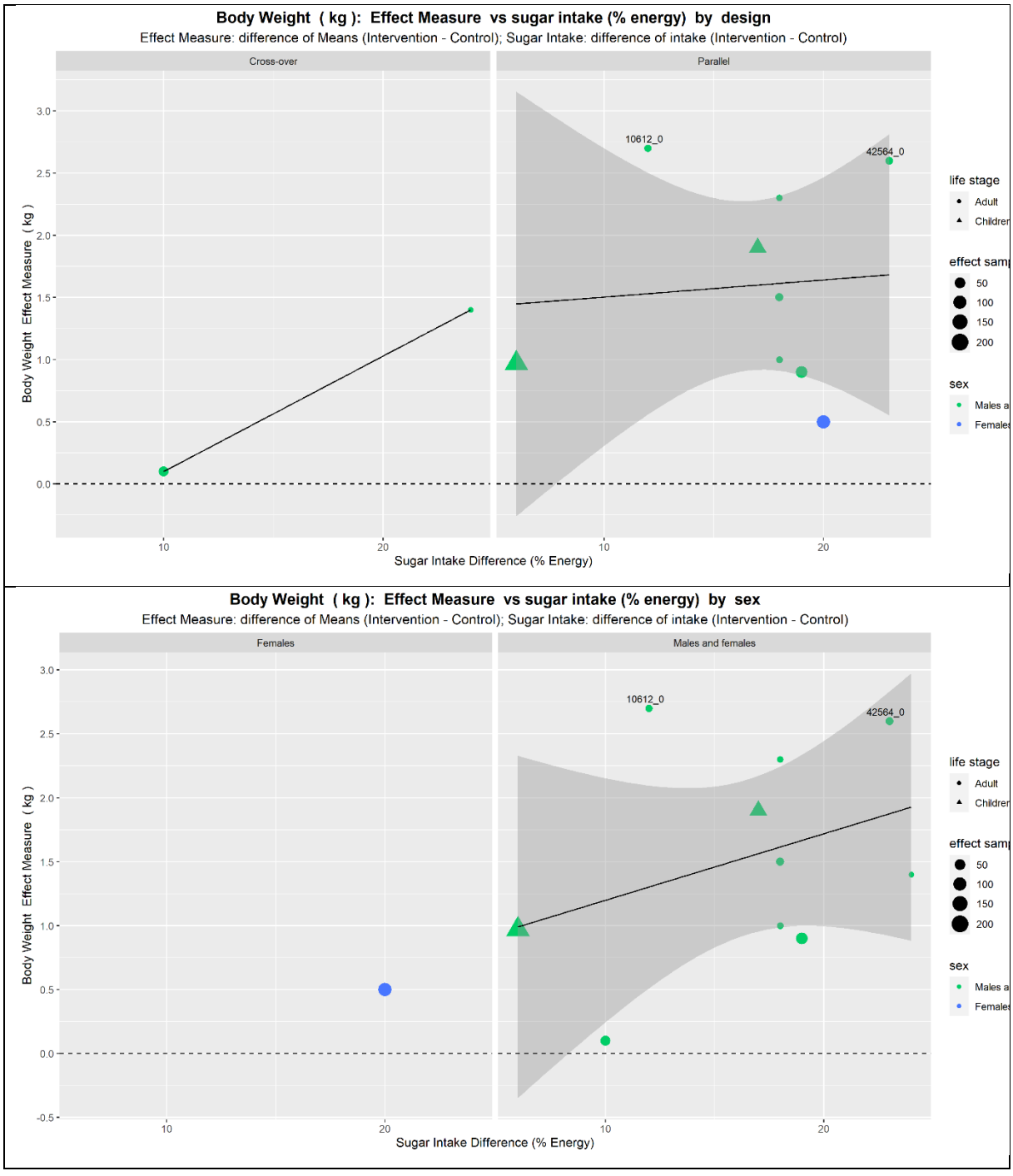
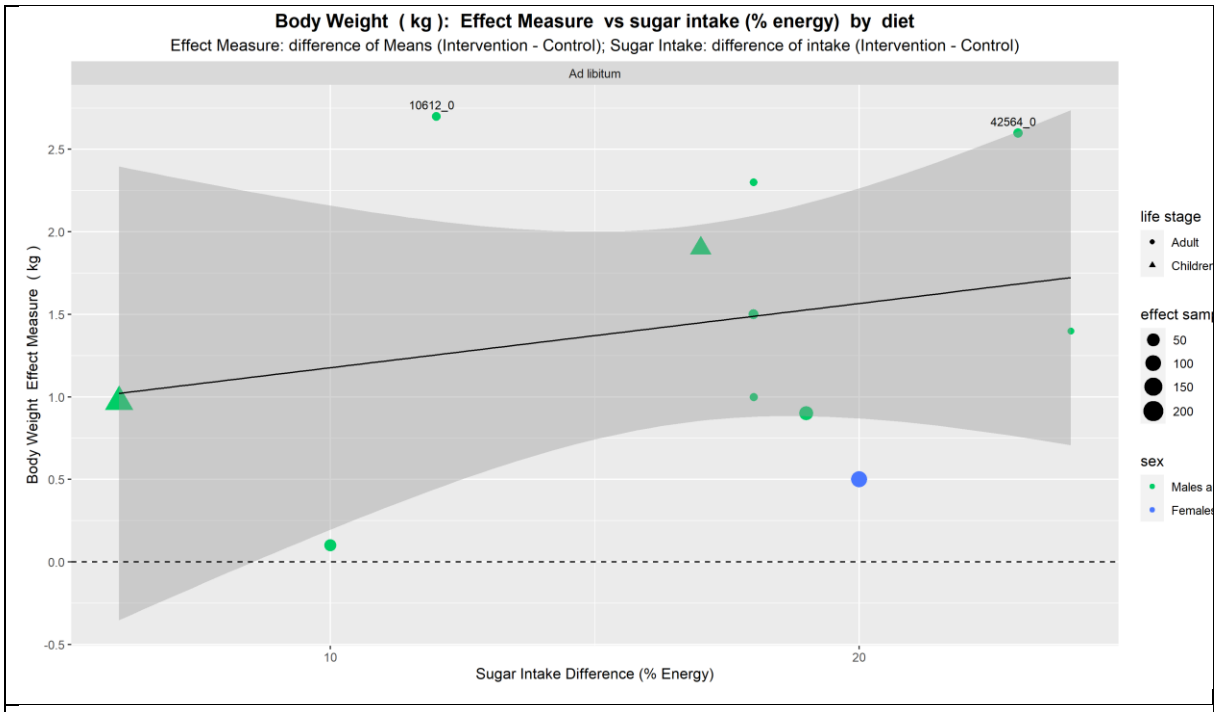


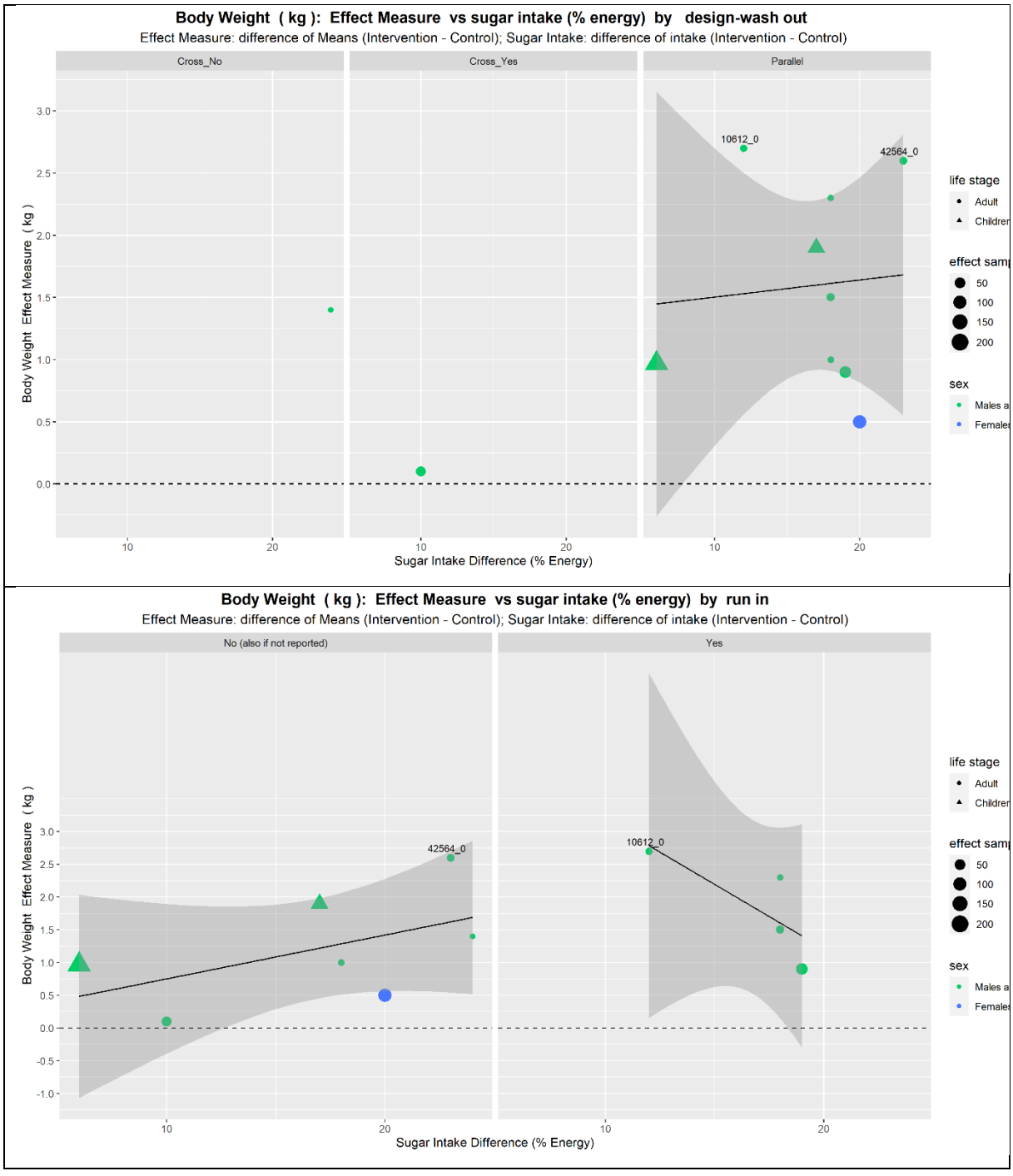
Figure 6: Scatterplots and linear regression between fasting insulin mean effect (Q1) and sugar intake arm difference (data not meta-analysed)

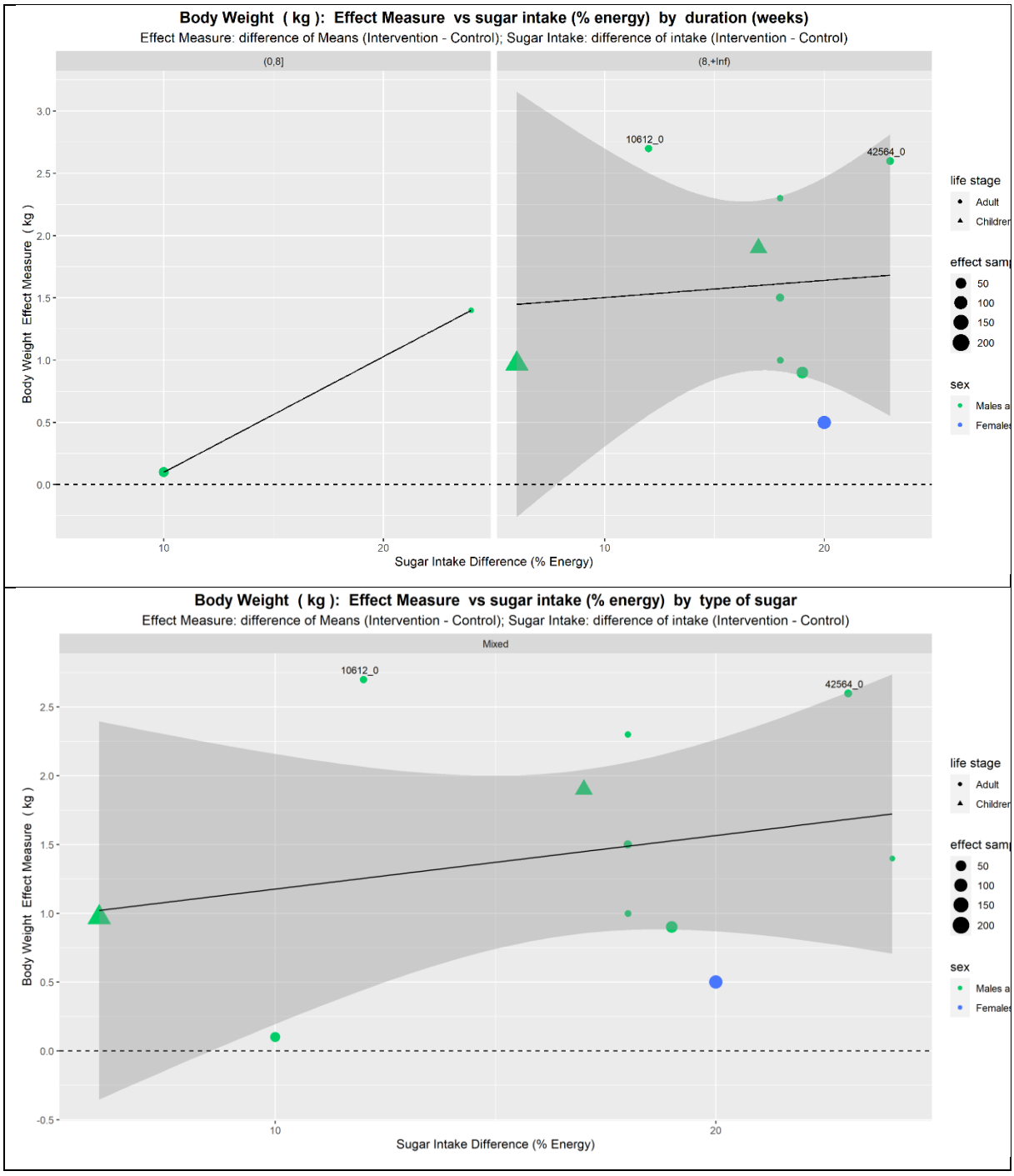
1.3.5. Body weight


Body Weight (kg): Effect Measure vs sugar intake (% energy) by design
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross-over
Parallel
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
10612_0
42564_0
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females
Body Weight (kg): Effect Measure vs sugar intake (% energy) by sex
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Females
Males and females
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
10612_0
42564_0
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females


Body Weight (kg): Effect Measure vs sugar intake (% energy) by diet
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Ad libitum
10612_0
42564_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females


Body Weight (kg): Effect Measure vs sugar intake (% energy) by source
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Beverages
Mix
10612_0
42564_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females


Body Weight (kg): Effect Measure vs sugar intake (% energy) by design-wash out
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Cross_No
Cross_Yes
Parallel
10612_0
42564_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females
Body Weight (kg): Effect Measure vs sugar intake (% energy) by run in
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
No (also if not reported)
Yes
42564_0
10612_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females


Body Weight (kg): Effect Measure vs sugar intake (% energy) by duration (weeks)
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
(0,8]
(8,+Inf)
10612_0
42564_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females
Body Weight (kg): Effect Measure vs sugar intake (% energy) by type of sugar
Effect Measure: difference of Means (Intervention - Control); Sugar Intake: difference of intake (Intervention - Control)
Mixed
10612_0
42564_0
Body Weight Effect Measure (kg)
Sugar Intake Difference (% Energy)
life stage
Adult
Children
effect sam
50
100
150
200
sex
Males a
Females

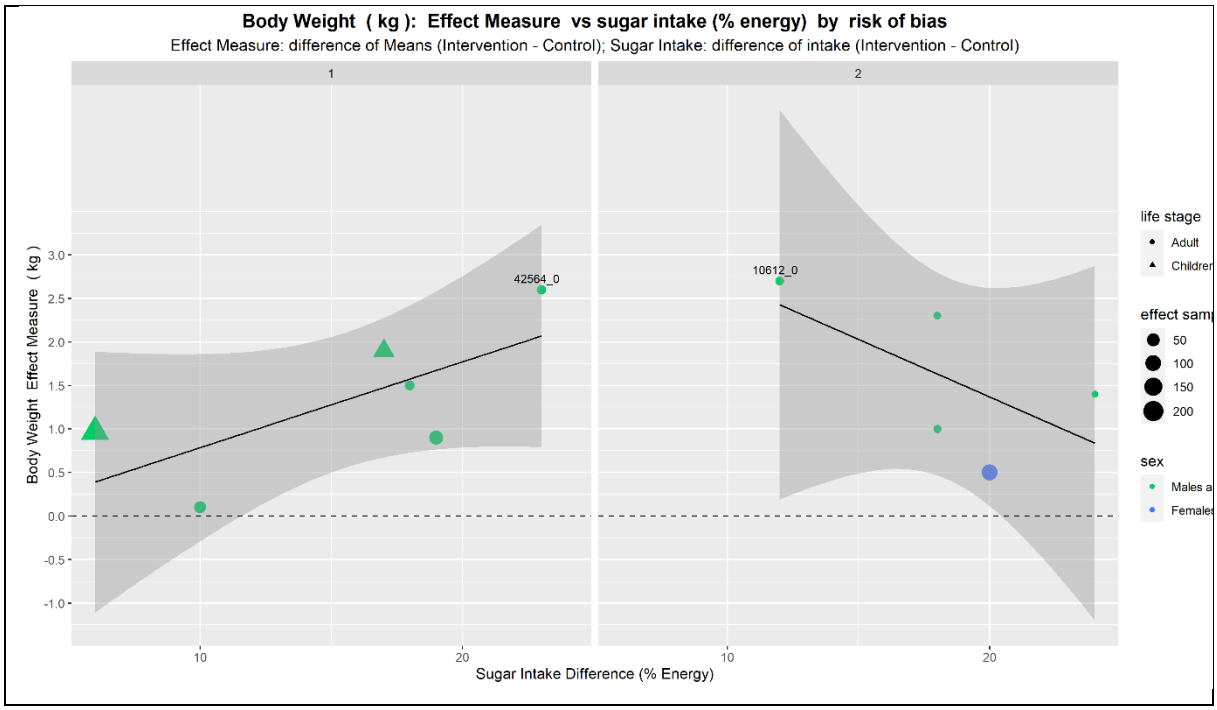


Figure 7: Scatterplots and linear regression between body weight mean effect (Q1) and sugar intake arm difference (data not meta-analysed)

1.4. Meta-analyses

Two types of meta-analyses were performed for both Q1 and Q2. Pooled mean effects and meta-regressive dose–response analysis, linear and non-linear. All the endpoints included in the assessment and all studies selected for relevance and not excluded in the pre-processing step were meta-analysed. For each endpoint, the pooled mean effect and its 95% CI and 95% PI are displayed in the forest plots using 0.82 as reference value for the correlation coefficient (Appendix L of the scientific opinion). The meta-regressive linear and non-linear dose–response analyses were used to investigate the relationship between the E% sugar dose arm difference as E% (Q1)/sugar E% from fructose compared with the same amount of glucose (Q2) and the effect on a subset of endpoints that included fasting triglycerides (Q1 and Q2), fasting glucose and fasting insulin, body weight (Q1) and uric acid (Q2). A summary of the meta-regressive dose–response analyses performed by endpoints and questions is displayed in Table 3.

Table 3: Meta-regressive dose–response models by endpoint and question

Outcome	Question	Linear	Non-linear
Triglycerides	Q1	X	X (non-linear coefficient not significant)
Fasting glucose	Q1	X	X (non-linear coefficient significant, but worst fitting)
Fasting insulin	Q1	X (model not sufficiently robust)	X (model not sufficiently robust)
Body weight	Q1	X (no dose–response – linear)	X (no dose–response – non-linear)
Triglycerides	Q2	X (no dose–response – linear)	X (no dose–response – non-linear)
Uric acid	Q2	X (no dose–response – linear)	X (too few observations for fitting a non-linear model)

To investigate the potential impact of individual studies and the methodological choices on the results of the meta-analyses (including dose–response), sensitivity analyses were performed looking at the extremes of the credible range for the correlation coefficient (0.5 and 0.99) compared with the best expert estimate of 0.82, and to the influence of individual studies using the one-at-a-time method on the regression parameters (only for dose–response meta-analyses).

A main uncertainty in the data was that the same E% from sugars in different studies could correspond to very different E% from sugars in the whole diet, depending on the contribution of the dietary fraction that was manipulated to total energy intake, and on the composition of the dietary fraction that was not manipulated. For most of the studies included, such information was not available. In addition, the target dose of sugars to be administered with the intervention, rather than the amount of sugars consumed (often not reported in studies conducted *ad libitum*) had to be used for data analysis since the achieved intake was not always available.

In this context, the only variable that could be investigated in relation to Q1 for different endpoints was the target (rather than the achieved) difference in sugars intakes between study arms.

The following assumptions were made:

- The dietary fraction that was not manipulated in the studies is comparable across arms regarding the macronutrient composition and, therefore, the sugar content both at baseline and at the end of the intervention.
- Between-arm differences in endpoint variables reflect the change that would occur in a group of individuals increasing sugars intake. This was effectively the case in studies in which the intervention aimed at increasing sugars intake, but not in studies in which the intervention aimed at reducing sugar intake.

1.4.1. Heterogeneity characterisation

The heterogeneity across the studies included in the BoE was first measured and then investigated with the scope of assessing whether specific methodological or clinical aspects could explain variability in the results.

The heterogeneity across studies addressing Q1, measured by the I^2 value (Higgins and Thompson, 2002), was between moderate and high according to the Cochrane classification (Higgins et al., 2020) for almost all the endpoints except those measuring body fatness (i.e. body fat, body weight, BMI, waist circumference). Heterogeneity across studies answering Q2 was also moderate or high for most of the endpoints (ectopic fat: liver fat and VAT; blood pressure: diastolic and systolic, fasting glucose and insulin, fasting triglycerides, total cholesterol and uric acid). Low variability across studies was only observed for HDL and LDL cholesterol concentrations. For the remaining endpoints the number of studies was very limited, making the measure of heterogeneity not meaningful.

The stratification of the studies by type of diet (i.e. isocaloric with neutral energy balance, isocaloric with positive energy balance, *ad libitum*) and type of sugars source (i.e. solid foods, beverages, mixtures of foods and beverages) did not help reducing heterogeneity except when few trials belonged to a single category. The latter was considered a methodological artefact and ignored. Despite the moderate to high unexplained heterogeneity across studies, the working group (WG) decided to pool the mean effects. This choice was made to avoid a narrative synthesis of the results.

1.4.2. Pooled estimates

Due to the large variability across studies it was decided to use a random effects model for the pooled estimates (Borenstein et al., 2009). Indeed the assumption that the trials are a random sample of a population of studies was considered the most realistic. The REstricted Maximum Likelihood method (REML) was used to estimate variability across studies. The Hakn correction was not used. The 95% CI and 95% PI values were generated for the pooled estimates to account for the sampling uncertainty in the prediction of the pooled mean effect (CI) and of the mean effect of a theoretical new study extracted from the same population (PI). The pooled mean effect estimates, their 95% CIs and 95% PIs are reported in the forest plots (Appendix G of the opinion).

A sensitivity analysis was conducted to explore the influence of the correlation coefficient value (0.82) on the pooled effect. The individual studies mean effect and the pooled mean effect were computed for values at the extreme of the credible interval 0.50 and 0.99 along with their 95% CIs. For individual studies, the influence of the correlation coefficient was assessed looking at any changes in the statistical significance of the mean effect. Since decreasing the correlation coefficient (from 0.82 to 0.5) could

lead only to larger 95% CIs, moving from significance to non-significance was the only expected effect (indicated with '1' in the forest plot under the heading 'r0.5'). The opposite is expected when increasing the correlation coefficient value from 0.82 to 0.99 (indicated with '2' in the forest plot under the heading 'r0.99'). No change in the statistical significance was indicated with '0' in the forest plot. The impact of the choice of the correlation coefficient value on the pooled mean effect could not be predicted upfront. It was assessed for each study and endpoint to understand whether magnitude and precision of the pooled effect was influenced by this methodological choice.

1.5. Meta-regressive dose–response analysis – linear relationship

A one-stage multivariate dose–response meta-analyses (Crippa et al., 2018) was set up to investigate the relationship between:

- sugar dose difference (E%) between arms and mean effect (i.e. mean difference or mean change difference between arms) – Q1. The analysis was performed for fasting triglycerides, fasting glucose and insulin and body weight;
- fructose intake versus the same amount of glucose intake (E%) and mean effect (i.e. mean difference or mean change difference between arms) – Q2. The analysis was performed for fasting triglycerides and uric acid.

In the one-stage approach, the effects of the studies are directly meta-analysed to derive a pooled dose–response curve. The two-stage approach requires fitting a dose–response curve to each study before meta-analysing them. Compared with the two-stage approach, the one-stage method has proven to be more efficient, as it allowed studies with less than three doses to be kept in the pooled model. Those studies represent the majority in our BoE.

First a linear shape was assumed. The models included as fixed effect the E% sugar dose difference between arms (Q1)/E% fructose dose vs the same amount of glucose (Q2) and, when appropriate, additional factors that were assessed for their potential to explain part of the variability of the effect across studies and their ability to modify the dose–response relationship (i.e. modifiers). A preliminary list of factors was suggested by the experts of the WG on sugars based on the biological plausibility of their influence. These were screened by visual inspection of the scatterplots showing dose–response stratified by factors' categories (Section 1.3). Continuous factors were discretised and used as categorical (body weight change difference, duration). Only those leading to statistically significant parameters were retained for the final models.

Mimicking the choice carried out for the pooled mean effect that would represent a meta-regressive model with only the intercept, a random factor was included to account for the variability explained by the study effect. Some studies reported results separately by sex. When randomisation was not carried out separately for males and females, it was concluded that results by sex could not be considered independent. Therefore, an additional random component was applied to reflect the hierarchical structure of the data. This occurred for fasting triglycerides (both Q1 and Q2) and fasting insulin (Q1). A compound symmetry structure was adopted to describe the correlation within studies. The classical assumptions of normality of residuals and homoscedasticity were made. They were investigated using graphical analysis of the residuals (empirical distribution of the residuals, scatter of the standardised residuals by dose and response), Q-Q plots (Wilk and Gnanadesikan, 1968) and formal testing using the Shapiro–Wilk test (Shapiro and Wilk, 1965).

The choice of the best model for each endpoint was based on three criteria considered concurrently: 1. goodness of fit assessed using the Akaike Information Criteria (AIC; Akaike, 1974); 2. significance of the parameters; 3. explained heterogeneity. Once the best model was identified, a sensitivity analysis was performed to check the robustness of the model with respect to the influence of individual studies using the one-at-a-time (Saltelli et al., 2009) leave out analysis. This method allows quantifying the change in the model parameter estimates when one study at a time is taken out from the BoE. Studies having a larger influence on the parameter estimates were inspected for potential errors and for peculiarities in the study population and/or the study settings. The WG discussed whether it was more appropriate to keep or to exclude the study from the BoE. When excluded, motivations are provided in this report. The presence of leverage and outliers studies was also assessed using the Hat-value (Belsley et al., 2005) with a threshold put at twice the mean of the hat values, and the Cook distance (Cook,

1977) with a threshold put at five times the mean of the Cook values. An additional sensitivity analysis was deemed useful to investigate the influence of the correlation coefficient on the parameter estimates given that its credible range was relatively large (from 0.5 to 0.99).

1.5.1. Meta-regressive linear dose–response analysis: fasting triglycerides (Q1)

A meta-regressive mixed-effects dose–response model was used to investigate the relationship between the E% sugar dose difference between arms and the fasting triglycerides (TG) mean effect (i.e. mean difference or mean change difference between arms). In total, 29 observations were eligible for the analysis. Among the factors identified by the experts as potential modifiers of the effect or of the dose–response displayed in the scatterplots (see section 1.3.1), the following were tested for inclusion in the model: type of sugar, source of sugar, type of diet, design combined with wash-out, RoB tier. Although the regression parameter for the factor reflecting the design combined with the wash-out was the only statistically significant, it was not retained in the model because it was largely influenced by a single study. Two random effects were included to account for the hierarchical structure of the data due to the study effect and the effect of sex within a single study that made the results for males and females (not randomised independently) correlated.

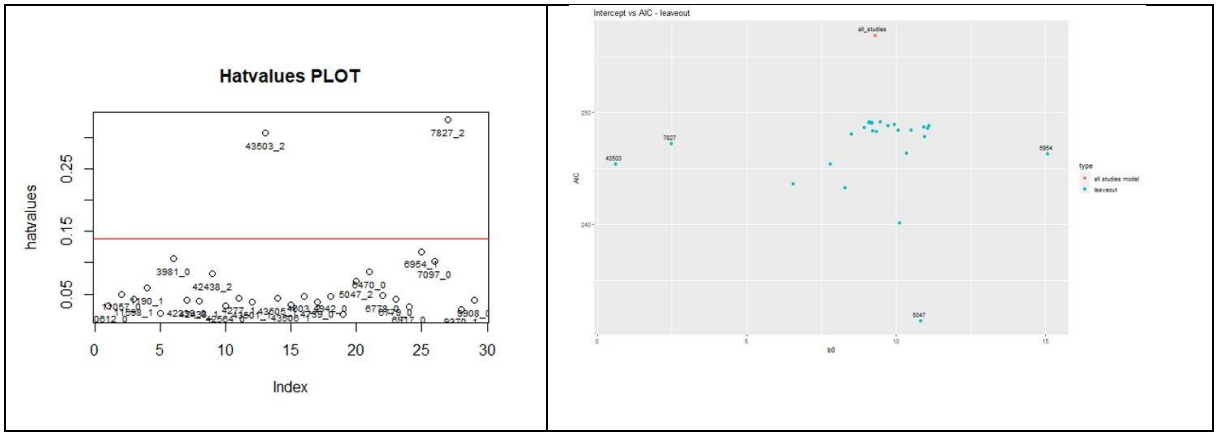
The equation for the final dose–response model is:

$$\boldsymbol{\theta_{jk} = \beta_0 + \beta_1 X + \epsilon_j + \rho_k} \qquad \text{Equation 22}$$

where

- θ_{jk} is the effect on fasting triglycerides observed in the j-th study and k-th sex
- X is the fixed effect due to the E% sugar dose difference between arms
- ϵ_j is the random effect due to the individual study
- ρ_k is the random effect due to sex.

Several diagnostics, the Hat indicator, the Cook distance and the influence analysis (one-at-a-time leave out analysis), identified one study (Moser et al., 1986), conducted on two subgroups of young women taking/not taking contraceptives, as highly influential because of the high sugars dose and the particularly small size of the effect (Figure 8). Since the results of the study were counter-conservative (i.e. very low responses at high doses), and their impact was to flatten the dose–response, it was decided to exclude the two observations from the dose–response analysis.



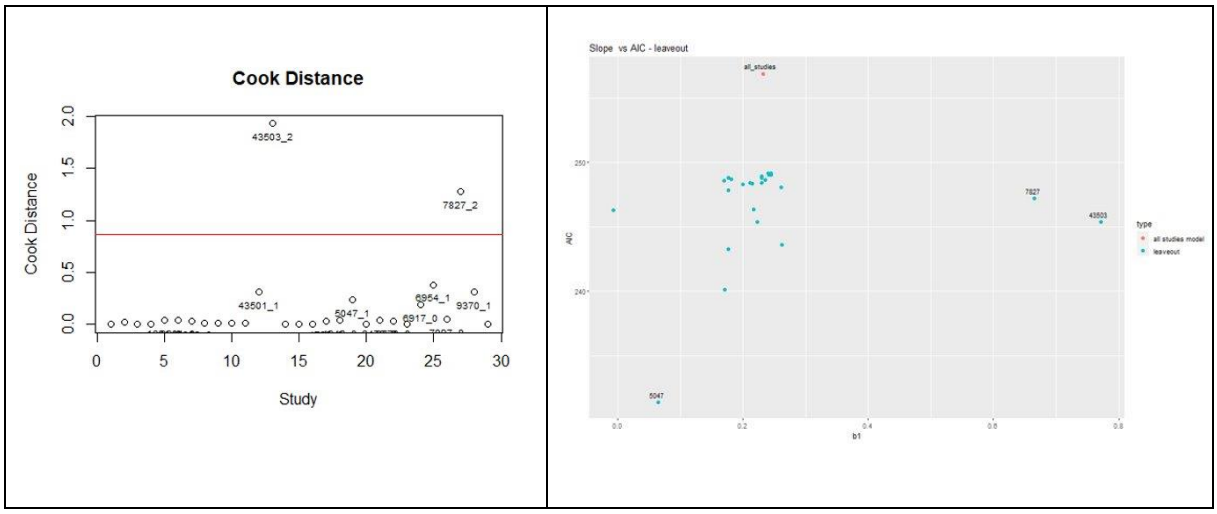


Figure 8: Model diagnostics – influence of individual studies on dose–response for TG

Therefore, the final model was set up on 27 observations with the sugars dose ranging from 6% to 30% E%.

The visual inspection of the Q–Q plot and the Shapiro–Wilk test provided indication for some deviation from normality of the residuals due to right skewness of the distribution of the effects. The log-transformation did not improved normality. Therefore, the original scale was maintained.

The heterogeneity explained by the sugars dose was limited (Cochran Q-test for modifier = 16.17). Therefore the residual heterogeneity remained high (Cochran Q-test for residual heterogeneity = 66.39) and statistically significant ($p < 0.0001$), indicating that other factors not identified in the BoE, or for which it was not possible to adjust due to the low number of studies available, play a role in explaining differences across studies. The final model indicates an expected increase in fasting TG of ~17 mg/dl (95% CI: 8.9, 25.8, $p < 0.01$) per each increase of 10 E% intake from sugars with a negative estimate of the intercept (−16.70 mg/dL, 95% CI: −32.88, −0.53, $p = 0.04$). In the final linear model, between-arm differences in sugars intake (E%) only accounted for ~20% of the variability across studies, therefore leaving most of the heterogeneity unexplained.

Results of model estimates and the display of the dose–response relationship are provided respectively in Table 4 and Figure 9.

Table 4: Fasting triglycerides – linear dose–response meta-analysis, goodness of fit, fixed and random effects estimates, measure of heterogeneity ()

Model goodness of fit						
logLik	**Deviance**	**AIC**	**BIC**			
−110.19	220.38	228.38	233.26			
Random effects						
	estimate	**Sqrt**	**n. levels**			
Study	0	0.0013	25			
Sex	106.54	10.32	27			
Test for residual heterogeneity			**Test for modifiers**			
QE	**df**	**p-val**	**QE**	**df**	**p-val**	
66.39	25	<0.0001	16.17	1	<0.0001	
Model results						
	estimate	**se**	**zval**	**p-val**	**CI.lb**	**CI.ub**
Intercept	−16.70	8.25	−2.02	0.043	−32.88	−0.53
dose	1.7325	0.4309	4.0208	<0.0001	0.89	2.58

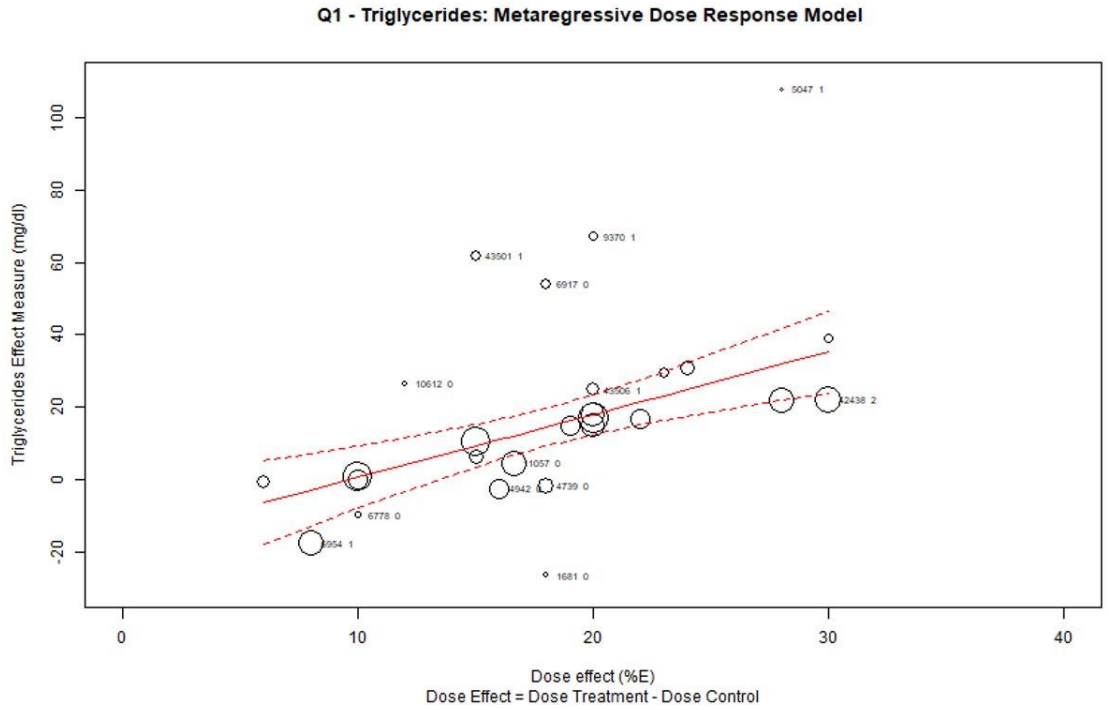


Figure 9: Meta-regressive linear dose–response model between the intake of added and free sugars (E%) and mean effect on fasting triglycerides

As illustrated in Figure 9 the effects of some studies are not fitted well by the model. This is probably due to the characteristics of the related study subjects being either hyper-insulinemic, hyper-triglyceridemic or overweight/obese. This consideration combined with the high heterogeneity unexplained by the model led the WG to conclude that this analysis could be used to conclude on the shape and direction of the dose–response relationship, but not to make a quantitative prediction of the effect of added or free sugars on fasting levels of triglycerides.

The sensitivity analysis performed to assess the influence of the correlation coefficient on the model parameters did not highlight a large variation of the estimates with respect to the range of its plausible values. The intercept estimate ranges between −20.5 and −16 for values of the correlation coefficient between 0.5 and 0.99 (−16.8 at 0.82) and the slope between 1.55 and 2.2 (1.7 at 0.82), as illustrated in Figure 10.

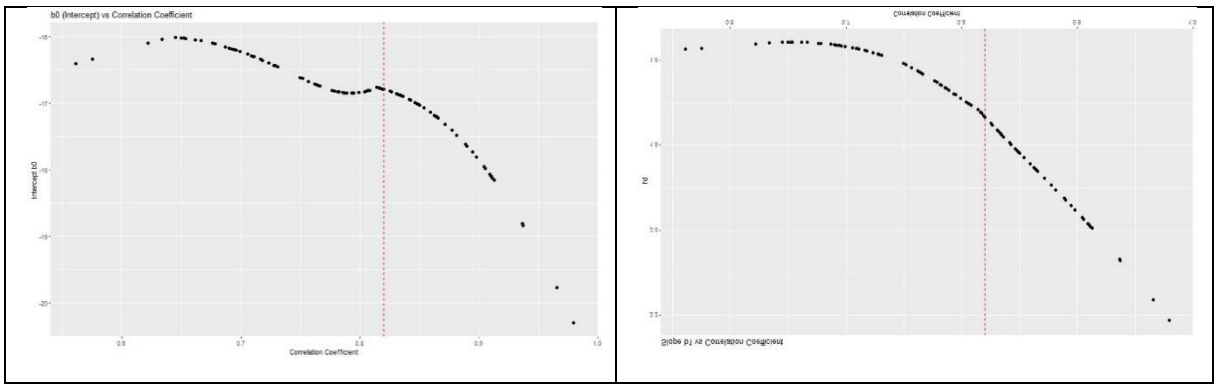


Figure 10: Influence analysis of the correlation coefficient on the model intercept (left) and slope (right)

1.5.2. Meta-regressive linear dose–response analysis: fasting triglycerides (Q2)

A meta-regressive mixed-effects dose–response model was also used to investigate the relationship between fructose versus the same amount of glucose and TG mean effect (i.e. mean difference or mean change difference between arms). Nine observations from eight studies were included in the analysis.

Although the model including the dose of fructose and glucose (E%) as explanatory variable fits data better than the null model (i.e. model with no explanatory variables), the CI of the regression coefficient estimate overlaps zero (95% CI: −3.06, 2.66). In the null model the overall mean effect is equal to 6.63 mg/dl, with a large sampling uncertainty expressed by a 95% CI (−2.85, 16.1) that includes zero. Therefore, it is not possible to conclude towards a dose–response relationship between the intake of fructose (vs glucose, E%) and the difference in fasting triglycerides (or difference of change from baseline).

1.5.3. Meta-regressive linear dose–response analysis on fasting glucose (Q1)

A meta-regressive mixed-effects dose–response model was used to investigate the relationship between the E% sugar dose difference between arms and the fasting glucose (FG) mean effect (i.e. mean difference or mean change difference between arms) expressed in mg/dl. In total, 19 observations from 18 studies were eligible for the analysis. Among the factors identified by the experts as potential modifiers of the effect or of the dose–response and displayed in the scatterplots (see Section 1.3.3), the following were tested for inclusion in the model: source of sugar, type of diet , design combined with wash-out, intervention duration, run-in, RoB tier. Although the model including factors related to the RoB and presence of run-in was providing the best relative fit in terms of the AIC criterion, the second explanatory variable was not retained in the final model since the significance of the associated regression parameter was largely influenced by a single study. The study was included as random effect.

The equation for the final dose–response model is provided below:

$$\boldsymbol{\theta_j} = (\boldsymbol{\beta_{01}} + \boldsymbol{\beta_{02}}\boldsymbol{X1}) + \boldsymbol{\beta_1}\boldsymbol{X2} + \boldsymbol{\epsilon_j} \qquad \text{Equation 23}$$

where

- θ_j is the mean effect on FG observed in the j-th study
- X1 is the differential fixed effect on the intercept due to the RoB2 with respect to RoB1 (X1 = 0 for RoB1, X1 = 1 for RoB2)
- X2 is the fixed effect due to the E% sugar dose difference between arms
- ϵ_j is the random effect due to the individual study.

Several diagnostics, the Hat indicator, the Cook distance and the influence analysis (one-at-a-time leave out analysis), identified one study (Moser et al., 1986), conducted on the subgroup of young women taking contraceptives, as highly influential (Figure 11) because of the high sugars dose and the particularly small size of the effect. Since the results of the study subgroup were counter-conservative (i.e. very low responses at high doses), and their impact was to flatten the dose–response, it was decided to exclude the observation from the dose–response analysis. Despite not being influential and showing a pattern fitting well the model, also the other subgroup (women not taking contraceptives) was dropped from the analysis since it came from the same study with randomisation performed for the two subgroups combined. Therefore, the final dose–response model was set up on 17 observations from 17 studies with E% sugar intake (difference between arms) ranging between 8% and 28%.

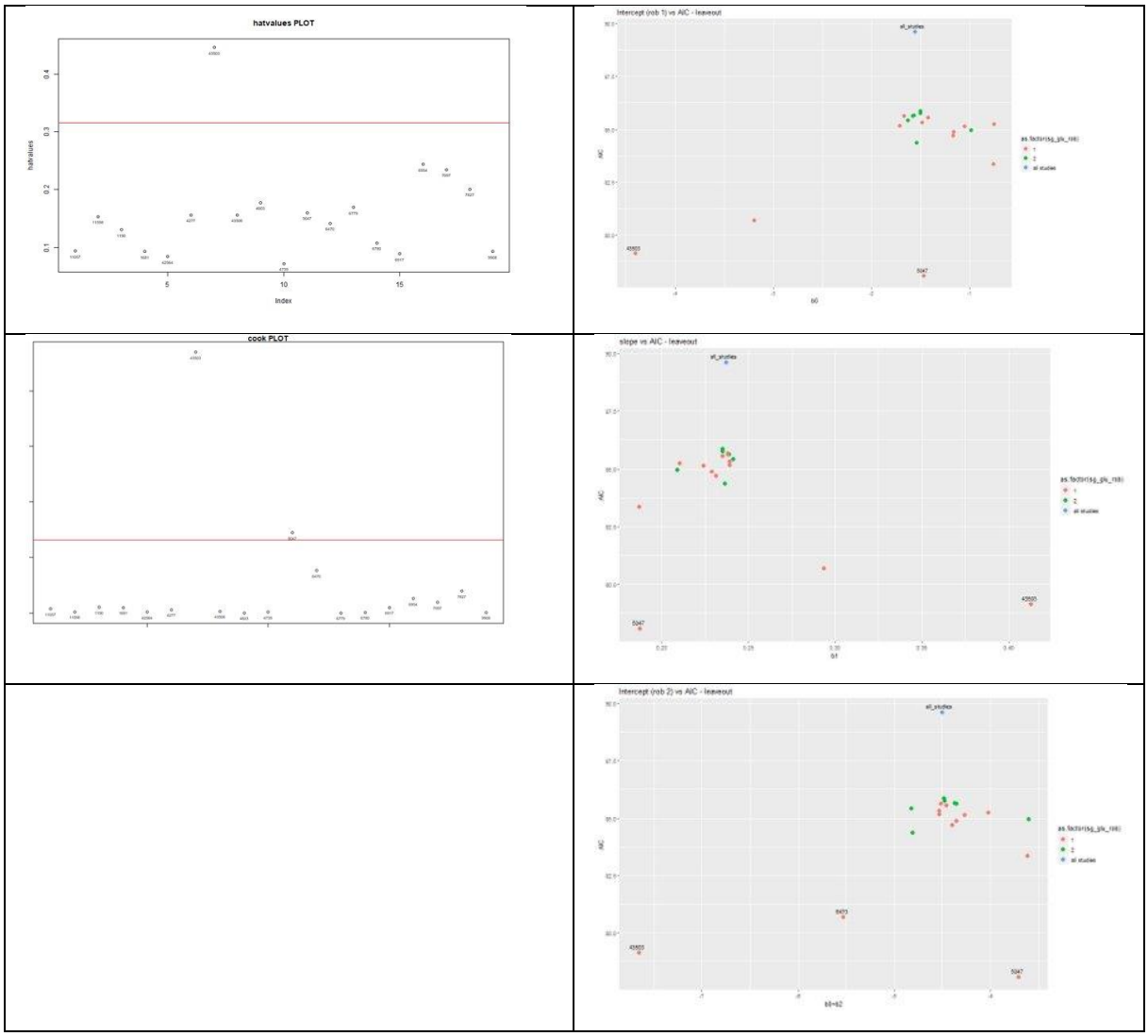


Figure 11: Model diagnostics – influence of individual studies on dose–response for FG

Residual heterogeneity remains high (Cochran Q-test = 43.26) and statistically significant ($p < 0.0001$) for the best fitting model, suggesting that other factors not identified in the BoE, or for which it was not possible to adjust due to the low number of studies, might play a role in explaining differences across studies. The model indicates an expected increase of ~4 mg/dl (95% CI: 1.7, 6.3; $p < 0.01$) of blood FG levels per each increase of 10 E% intake from sugar. Adjusting for RoB leads to higher absolute FG mean expected levels for the same level of sugars intake when considering RCTs at low RoB (tier 1; intercept = −4.2 mg/dl; 95% CI: −8.4, 0.03) compared with RCTs at moderate RoB (tier 2; intercept = −7.4; 95% CI: −13.91, −0.95). Between-arm differences in sugars intake (E%) and RoB only accounted for 25.6% of the variability across studies, therefore leaving most of the heterogeneity unexplained. Therefore, it was deemed appropriate to use this analysis to conclude on the shape and direction of the dose–response relationship, but not to make a quantitative prediction of the effect of sugars on fasting levels of triglycerides.

Results of model estimates and the display of the dose–response relationship are provided in Table 5 and, Figure 12, respectively.

Table 5: Meta-regressive dose–response analysis – fixed and random effects estimates, measure of heterogeneity (fasting glucose)

Model goodness of fit						
logLik	**Deviance**	**AIC**	**BIC**			
−33.44	66.87	74.87	77.43			
Random effects						
	estimate	**sqrt**	**n. levels**			
study	3.8	1.95	17			
Test for residual heterogeneity			**Test for modifiers**			
QE	**df**	**p-val**	**QE**	**df**	**p-val**	
43.26	14	<0.0001	14.89	2	0.0006	
Model results						
	estimate	**se**	**zval**	**p-val**	**CI.lb**	**CI.ub**
Intercept (RoB1)	−4.2	2.16	−1.95	0.052	−8.44	0.03
Intercept (RoB2)	−3.23	1.15	−2.82	0.005	−5.48	−0.99
Dose	0.40	0.12	3.46	0.0005	0.17	0.63

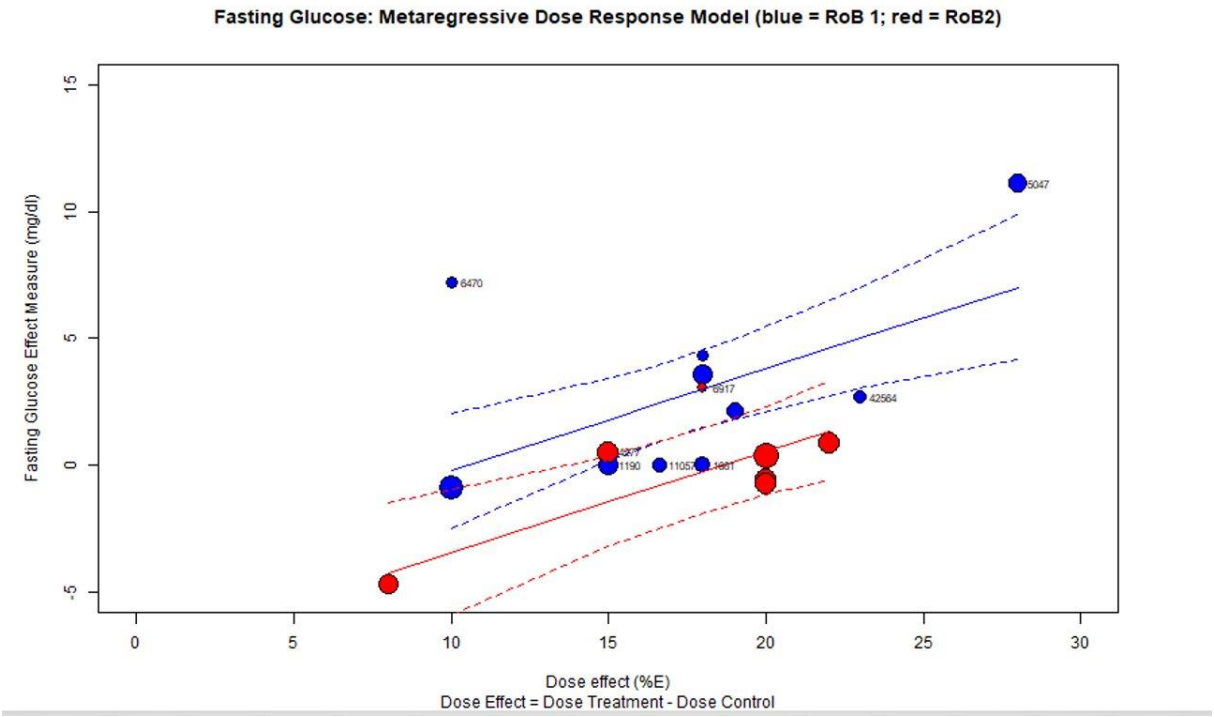


Figure 12: Meta-regressive linear dose–response model between the intake of added (and free) sugar (E%) and mean effect on fasting glucose

As illustrated in Figure 12, one study effect is not fitted well by the model. This is probably due to the characteristics of the related study subjects being overweighted/obese people.

The sensitivity analysis performed to assess the influence of the correlation coefficient on the model parameters did not highlight a large variation of the estimates with respect to this methodological choice.

The intercept estimate ranges approximately between −5.5 and −3.5 and −9 and −7 for values of the correlation coefficient between 0.5 and 0.99 (−4.5 and −8 at 0.82) respectively for RoB1 and RoB2. The slope estimate ranges approximately between 0.36 and 0.47 (0.42 at 0.82) as illustrated in Figure 13.

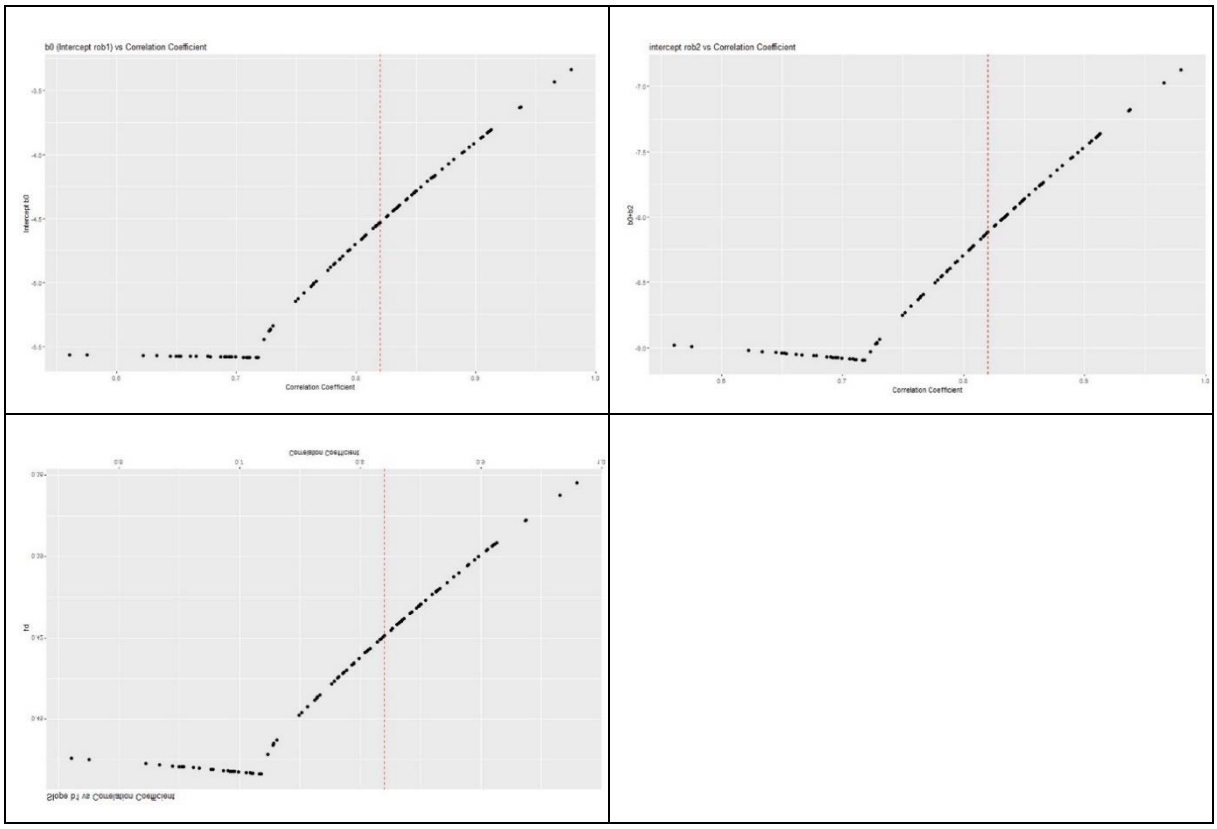


Figure 13: Influence analysis of the correlation coefficient on the model intercepts for RoB1 and RoB2 (above) and slope (below)

1.5.4. Meta-regressive linear dose–response analysis on fasting insulin (Q1)

A meta-regressive mixed-effects dose–response model was used to investigate the relationship between the E% sugar intake difference between arms and the fasting insulin (FI) mean effect (i.e. mean difference or mean change difference between arms). In total, 16 observations were eligible for the analysis. Among the factors identified by the experts as potential modifiers of the effect or of the dose–response and displayed in the scatterplots (refer to Section 1.3.4), the following were tested for inclusion in the model: type of diet, design combined with wash-out, duration of the intervention, presence of run-in, RoB tier. Although the regression parameters for the factor reflecting the design combined with the wash-out was statistically significant both as modifier of the effect and of the dose–response relationship and was providing a better fit to the data, it was not retained in the final model since it was largely influenced by a couple of studies. Two random effects were included to account for the hierarchical structure in the data due to the study effect and the effect of sex within a single study that made the results for the two subgroups (not randomised independently) correlated.

The equation for the final dose–response model is provided below:

$$\boldsymbol{\theta_{jk} = \beta_0 + \beta_1 X + \epsilon_j + \rho_k} \qquad \text{Equation 24}$$

where

- θ_{jk} is the effect on FI observed in the *j-th* study and sex *k-th*
- X is the fixed effect due to the E% sugar intake difference between arms

- ϵ_j is the random effect due to the individual study
- ρ_k is the random effect due to the sex.

Several diagnostics, the Hat indicator, the Cook distance and the influence analysis (one-at-a-time leave out analysis), identified one study (Moser et al., 1986), conducted on two subgroups of young women taking/not taking contraceptives, as highly influential (Figure 14) because of the high sugars intake and the particularly small size of the effect. Since the results of the study were counter-conservative (i.e. null response at very high intakes), and their impact was to flatten the dose–response, it was decided to exclude the two observations from the dose–response analysis. Therefore the final model was set up on 14 observations with a E% sugar intake (difference between arms) ranging between 10% and 28%.

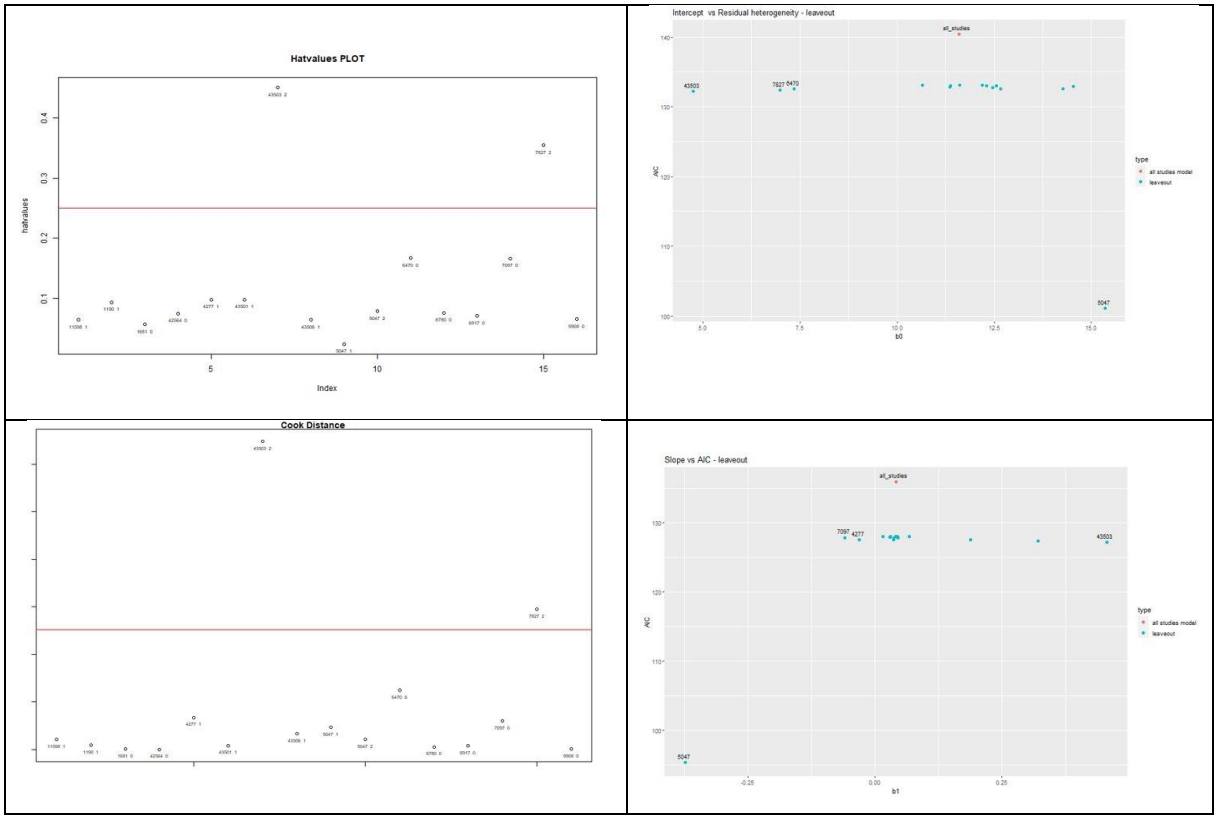


Figure 14: Model diagnostics – influence of individual studies on dose–response for FI

The visual inspection of the Q–Q plot and the Shapiro–Wilk test provided an indication for some deviation from normality of the residuals due to right skewness of the distribution of the effects particularly due to one study on hyper-insulinemic people. The log-transformation did not improved normality. Therefore, the original scale was maintained.

The heterogeneity explained by the intake was limited (Cochran Q-test for modifier = 3.67). Therefore the residual heterogeneity remained extremely high (Cochran Q-test for residual heterogeneity = 203.52) and statistically significant ($p < 0.0001$), indicating that other factors not identified in the BoE, or for which it was not possible to adjust due to the low number of studies available, played a role in the explanation, but accounted for less than 2% of the variability across studies, therefore leaving almost all of the heterogeneity unexplained. In addition the model appeared highly sensitive to the influence of a couple of studies Israel et al. (1983) and Lewis et al. (2013) and to other methodological choices (i.e. hypothesised level of the correlation between observations at the beginning and the end of the intervention in the parallel trials, and across intervention groups in the cross-over trials) as illustrated in Figure 15. Therefore, the model was not considered sufficiently robust to be used to draw conclusions on the shape and the strength of the dose–response relationship.

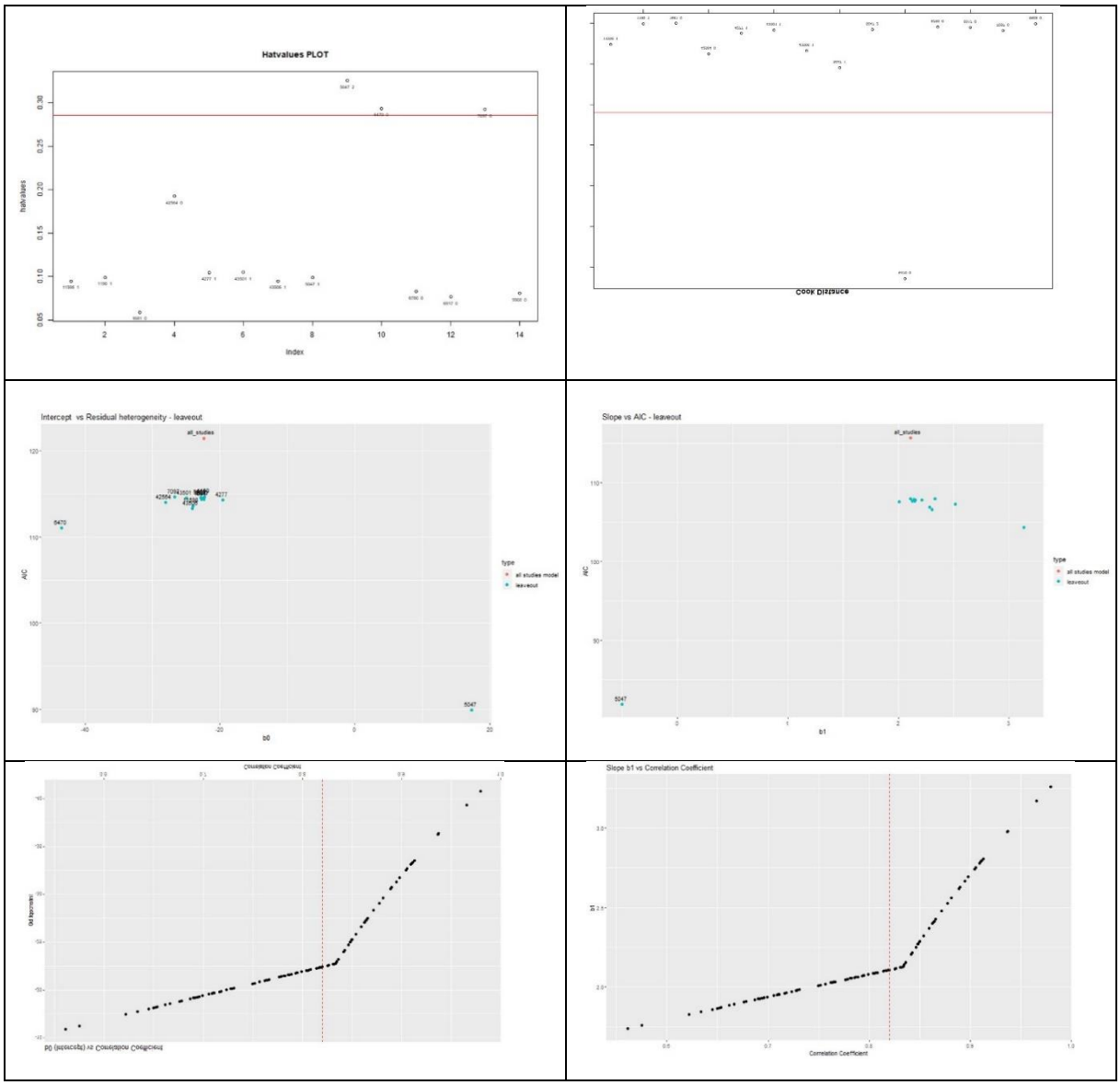


Figure 15: Model diagnostics – influence of individual studies and of the correlation coefficient on dose–response for FI

1.5.5. Meta-regressive linear dose–response analysis on body weight change (Q1)

A meta-regressive mixed-effects dose–response model was performed to investigate the relationship between the E% sugar intake difference between arms and the body weight (BW) change mean effect (i.e. mean difference or mean change difference between arms). In total, 11 observations were eligible for the analysis. The intake of sugars expressed as E% could not significantly explain the variability in the between-arm differences in BW changes (Cochran Q-test for modifier = 0.73 and Cochran Q-test for residual heterogeneity = 14.02). In fact, the fit of the model measured by the AIC, equal to 36.1, was not dissimilar to that of the model with no explanatory variables (AIC equal to 36.5). In addition, the regression parameter expressing the effect of the sugar intake on the BW change was extremely low and its CI overlapped zero (0.05; 95% CI: −0.0623, 0.1582; p = 0.3941). Due to the previous considerations, the impact of other variables as possible modifiers of the effect was not explored.

Therefore, evidence does not support a linear dose–response relationship between the intake of sugars as E% *ad libitum* and body weight change despite the pooled mean estimated BW change being equal to 1.15 kg and different from zero, also taking into consideration the sampling uncertainty (95% CI: 0.53, 1.77).

1.5.6. Meta-regressive linear dose–response analysis on uric acid (Q2)

A mixed-effects dose–response meta-regressive model was also used to investigate the relationship between the E% sugar intake from fructose versus the same amount of E% sugar from glucose and the uric acid (UA) mean effect (i.e. mean difference or mean change difference between arms). The study effect was treated as random. Only five observations were eligible for the analysis. Lack of data did not allow investigation of the dose–response relationship and estimation of the regression parameters with sufficient precision. In fact, the intake of fructose versus glucose expressed as E% could not significantly explain the variability in the between-arm differences in the UA (Cochran Q-test for modifier = 1.63 and Cochran Q-test for residual heterogeneity = 8.46). The fit of the model measured by the AIC, equal to 7.05, was larger than that for the model with no explanatory variables (AIC equal to 4.7) showing a poor relative fit.

In the null model the overall mean effect is equal to 0.12 mg/dl with a large sampling uncertainty expressed by a 95% CI (−0.16, 0.40) that includes zero. Therefore it is not possible to conclude a dose–response relationship between difference in UA (or difference of change from baseline) and E% fructose vs glucose. It is not possible also to conclude if there was an overall effect, due to the high heterogeneity across studies and the large sampling uncertainty.

1.6. Mixed-effects dose–response meta-analysis – non-linear relationship

A non-linear shape was also used to fit the dose–response relationship. A spline is generally a convenient tool to characterise a non-linear dose–response relationship since it allowed a good fit of the data with low degree polynomials to be achieved. Some aspects are critical when setting a spline, particularly the degree of the polynomial, the number and location of the knots and the constraints applied to the function. Increasing the degree of the polynomial and the number of knots increases the goodness of fit but makes interpretation more difficult and increases the risk of overfitting. The Restricted Cubic Spline (RCS) with three knots (Desquilbet and Mariotti, 2010; Crippa et al., 2018) represents a good balance in this respect. The RCS is a piecewise polynomial function each piece of which is a cubic polynomial. The function is derived meeting some constraints: 1. It must equal true observations in the knots; 2. The function itself and its first two derivatives are continuous in the inner knots; 3. It is linear in the extremes (before the first knot and after the last knot). The advantage of using an RCS with three knots lies in the requirement that only three parameters need to be estimated, having a robust behaviour at the tails of the distribution, and allowing the expression of a non-linear relationship as a linear one using a convenient transformation of the predictor (Crippa et al., 2018). Location of the knots remains a critical issue that was addressed by resorting to sensitivity analysis. A one-stage approach was used. This choice allowed the possibility to keep all the studies in the analysis that otherwise would have excluded as those with less than three doses. Similar to the approach for the linear model, the intercept was kept in the model. The fixed effect included the E% sugar intake difference between arms (Q1) and the modifiers of the effect or of the dose–response relationship already identified for the linear model. Only those leading to parameters for which 95% CI values did not overlap zero were retained for the final models. Two random factors were included in the model to account for the hierarchical structure of the data that included studies and sex subgroups in some of them. A compound symmetry structure was adopted for the correlation within studies. The non-linear models' performance was assessed considering the level of residual heterogeneity (Q-Cochrane), the relative goodness of fit (AIC) and the sampling uncertainty in the parameter estimates (95% CI of the estimates). The WG decided to fit the non-linear model only for those endpoints for which a linear dose–response could be established (triglycerides and FG for Q1).

The two equations formalising the RCS are provided below:

$$\boldsymbol{\theta_{jk} = (\beta_{01} + \beta_{02} X_0) + \beta_1 X_1 + \beta_2 \widetilde{X_1} + \epsilon_j + \rho_k} \quad \text{Equation 25}$$

$$\widetilde{\boldsymbol{X}} = \frac{(\boldsymbol{x - k_1})_+^3 - \frac{\boldsymbol{k_3 - k_1}}{\boldsymbol{k_3 - k_2}} \cdot (\boldsymbol{x - k_2})_+^3 + \frac{\boldsymbol{k_2 - k_1}}{\boldsymbol{k_3 - k_2}} \cdot (\boldsymbol{x - k_3})_+^3}{(\boldsymbol{k_3 - k_2})^2} \quad \text{Equation 26}$$

where

- θ_{jk} is the effect on the endpoint observed in the *j-th* study and sex *k-th*
- $\boldsymbol{\beta_{01}}$ intercept (or intercept for the reference category of a modifier, if included)
- $\boldsymbol{\beta_{02}} X_0$ differential component of the intercept when a modifier is included
- X is the E% sugar intake difference between arms
- $\widetilde{\boldsymbol{X}}$ is a convenient transformation of the E% sugar intake difference between arms
- ϵ_j is the random effect due to the individual study
- ρ_k is the random effect due to the sex
- k_1, k_2, k_3 are the three knots
- $u_+ = \begin{cases} u & if\ u_+ > 0 \\ 0 & otherwise \end{cases}$.

The advantage of this formal expression of the RCS is that it allowed testing of whether a non-linear model fits the data better than a linear model, simply by testing whether $\boldsymbol{\beta_2}$ is significantly different from zero. In fact, the transformation $\tilde{X}$ comprises all the non-linear component of the model.

1.6.1. Meta-regressive non-linear dose–response analysis: triglycerides (Q1)

As described above, an RCS was used to investigate whether the relationship between the E% sugar intake difference between arms and the triglycerides (TG) mean effect (i.e. mean difference or mean change difference between arms) could be better described by a non-linear model. In total, 27 observations were eligible for the analysis (after exclusion of the two observations from Moser et al., 1986). Only the sugar intake was included as an explanatory variable to mimic the choice carried out for the linear model and to make the results of the linear and non-linear shape comparable. Two random effects were included to account for the two hierarchical levels in the data (i.e. studies and sex subgroups).

To investigate the influence of the knot location, the model was run for several choices giving results of AIC, Higgins I^2 and parameter estimates as described in Table 6. The knots location corresponds to different levels of sugar intake difference (E%).

Table 6: Triglycerides: non-linear model (RCS) estimates at different knots locations

Model	Knots location	AIC	Residual heterogeneity I^2	Intercept (95% CI)	Beta1 (95% CI)	Beta2 (95% CI)
1	10, 15, 20	222.51	60.9%	−28.95 (−60.73, 2.83)	2.79 (0.27, 5.30)	1.02 (−3.32, 1.27)
2	10, 20, 26	222.44	60.8%	−26.02 (−50.99, −1.05)	2.46 (0.73, 4.18)	−0.81 (−2.49, 0.87)
3	12, 18, 24	222.52	60.9	−25.62 (−51.05, −0.19)	2.42 (0.67, 4.18)	−0.90 (−2.88, 1.09)
4	15, 18, 21 (25th, 50th, 75th centile)	222.61	61.2%	−24.50 (−49.38, 0.39)	2.32 (0.65, 4.0)	−0.76 (−2.63, 1.10)

5	15, 20, 25	222.48	60.9%	−24.22 (−46.72, 1.72)	2.28 (0.84, 3.71)	−0.91 (−2.84, 1.02)
6	18, 22, 26	222.35	60.8%	−23.59 (−44.08, −3.10)	2.21 (0.96, 3.45)	−1.09 (−3.19, 1.01)
7	20, 23, 26	222.28	60.7%	−23.37 (−43.11, −3.64)	2.18 (1.0, 3.36)	−1.25 (−3.55, 1.06)
8 (centiles 5%, 50%, 95%)	8.6, 18, 29.4	222.41	60.9%	−26.95 (53.09, −0.82)	2.55 (0.68, 4.43)	−1.26 (−3.82, 1.30)

The eight model shapes and the related 95% CI of the predicted mean effects on fasting triglycerides are displayed in Figure 16.

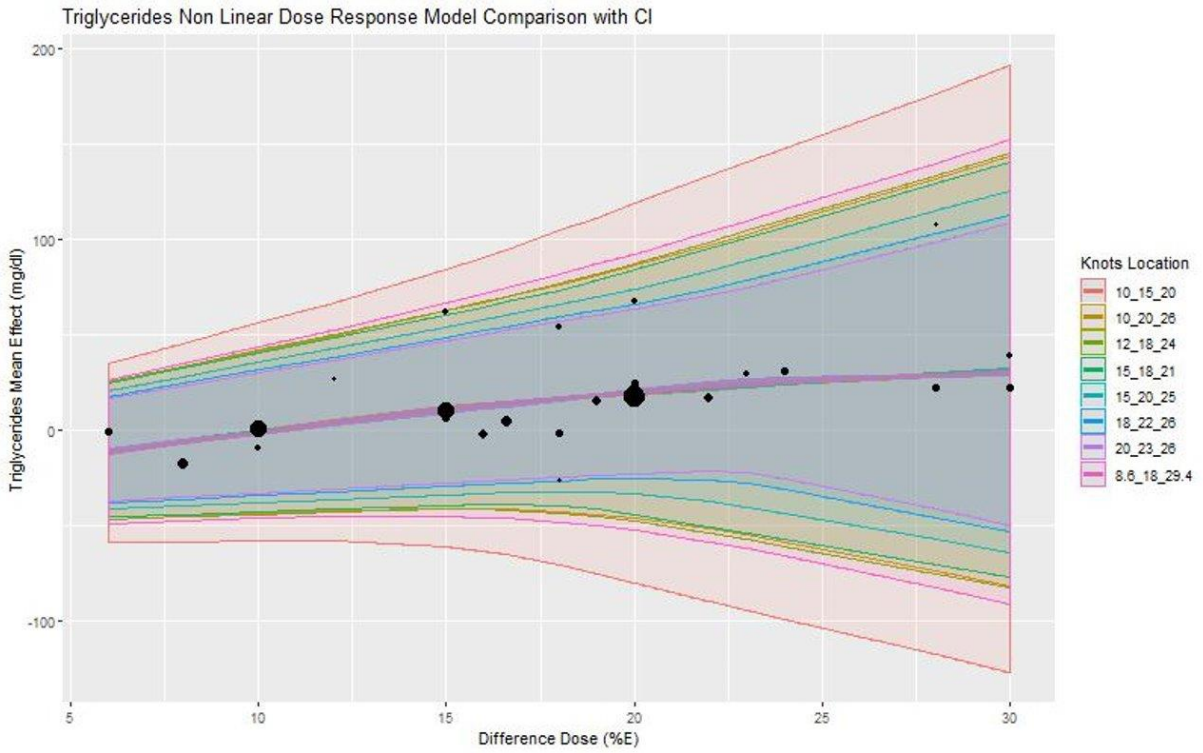


Figure 16: Fasting triglycerides – Comparison of non-linear dose–response models with different knot locations

For all the eight models the residual heterogeneity remained moderate (~60%) and the relative fit similar (AIC ~222). The 95% CI of the estimate of $\boldsymbol{\beta}_2$ always overlapped zero, indicating that a non-linear fit to the dose–response relationship is not justified.

1.6.2. Meta-regressive non-linear dose–response analysis: triglycerides (Q2)

An RCS was used to investigate whether the relationship between the E% sugar intake difference between arms and BW mean effect (i.e. mean difference or mean change difference between arms) could be better described by a non-linear model.

In total, nine observations were eligible for the analysis. The intake of sugars expressed as E% cannot significantly explain the variability in the between-arm differences in BW changes (Higgins I^2 for residual heterogeneity = 91.3%). In addition, the regression parameter expressing the non-linear effect of the sugar intake on the triglycerides change is low and its CI overlaps zero (0.5169; 95% CI: −6.6437,

7.6774). Due to the previous considerations, the impact of other variables as possible modifiers of the effect was not explored. Therefore, evidence does not support a non-linear dose–response relationship between the intake of sugars as E% *ad libitum* and triglycerides.

1.6.3. Meta-regressive non-linear dose–response analysis: fasting glucose (Q1)

An RCS was used to investigate whether the relationship between the E% sugar intake difference between arms and FG mean effect (i.e. mean difference or mean change difference between arms) could be better described by a non-linear model. In total, 17 observations were eligible for the analysis (after exclusion of the two observations from Moser et al., 1986). The E% sugar intake difference between arms and the RoB tier were included as fixed effects mimicking the choice carried out for the linear model. This made the results of the linear and non-linear shape comparable. Two random effects were included to account for the two hierarchical levels in the data (i.e. studies and sex subgroups). The compound symmetry structure was used to describe the correlation within studies.

To investigate the influence of the knots location, the model was run for three of those giving results of AIC, Higgins I^2 and parameter estimates as described in Table 7. The locations correspond to different levels of E% sugar intake difference.

Table 7: Fasting glucose: non-linear model (RCS) estimates at different knots locations

Model	Knots location	AIC	Residual het I^2	Intercept RoB1 (95% CI)	Intercept RoB2 (95% CI)	Beta1_RoB1 (95% CI)	Beta2_RoB2 (95% CI)
1	10, 20, 25	77.47	63.1%	0.37 (−2.17, 2.90)	−2.83 (−3.86, −1.81)	0.21 (0.01, 0.40)	0.33 (0.10, 0.56)
2	12, 16, 20	81.17	66.6%	0.61 (−2.46, 3.67)	-3.18 (-4.15, −2.20)	0.22 (−0.02, 0.46)	0.30 (0.02, 0.57)
3 (centiles 25%, 50%, 75%)	15, 18, 20	79.22	64.9%	0.07 (−2.45, 2.58)	−3.03 (−4.02, −2.04)	0.26 (0.08, 0.44)	0.27 (0.06, 0.48)

The three model curves and the related 95% CI of the predicted mean effects on FG when studies are classified in the RoB 1 and 2 tiers and are displayed in Figure 17.

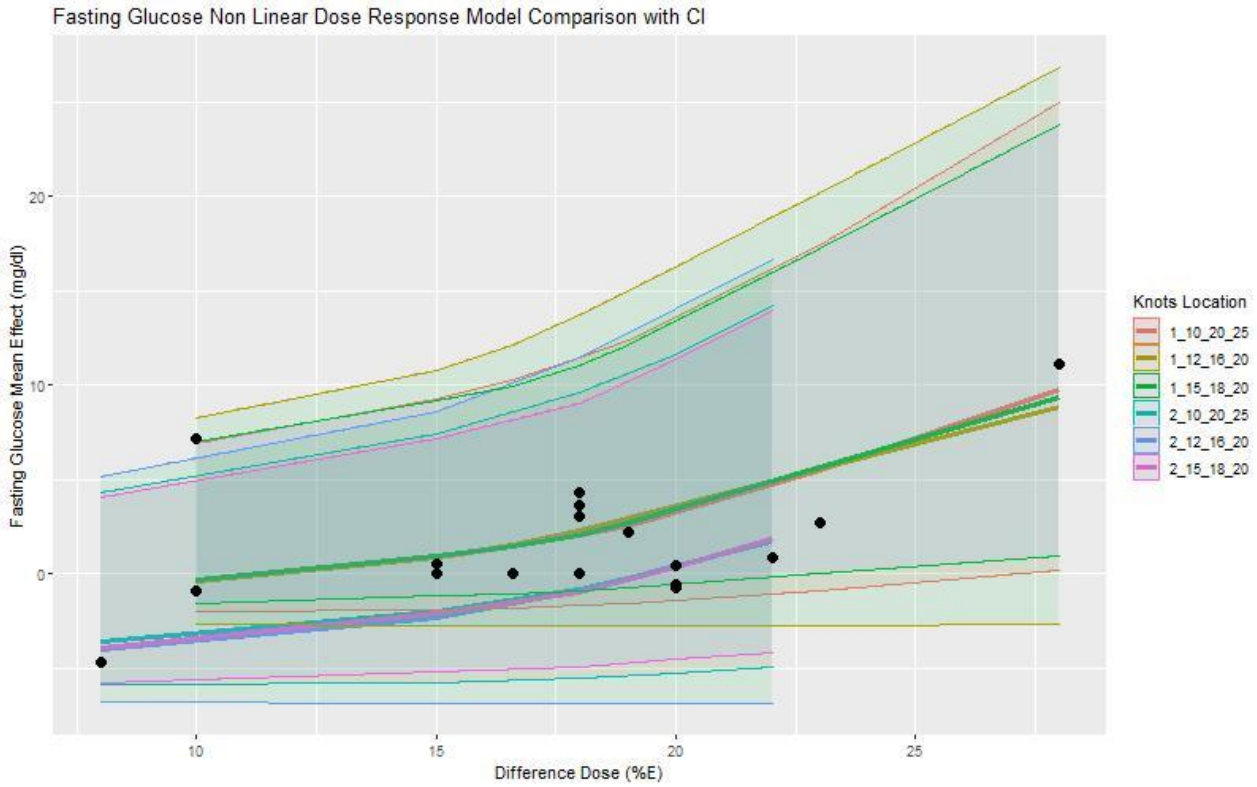


Figure 17: Fasting glucose – comparison of non-linear dose–response models with different knot locations

For all the three models the residual heterogeneity remained high (Higgins I^2 above 63%) and the relative fit was similar (AIC range between 77 and 81). The model with the best performance in terms of AIC, residual heterogeneity and precision of the mean effect estimates is model 1 (knots located at 10, 20, 25). A higher effect is predicted by studies with low RoB compared with those with medium RoB. The 95% CI of the estimate of $\boldsymbol{\beta_2}$ did not overlap zero for any of the four models (except marginally for model 2 and RoB = 1 subgroup) suggesting that a non-linear fit to the dose–response relationship might be justified. However due to the better fit of the linear model (AIC 74.87 versus 77.47), the latter was retained for drawing conclusions.

1.6.4. Meta-regressive non-linear dose–response analysis: fasting insulin (Q1)

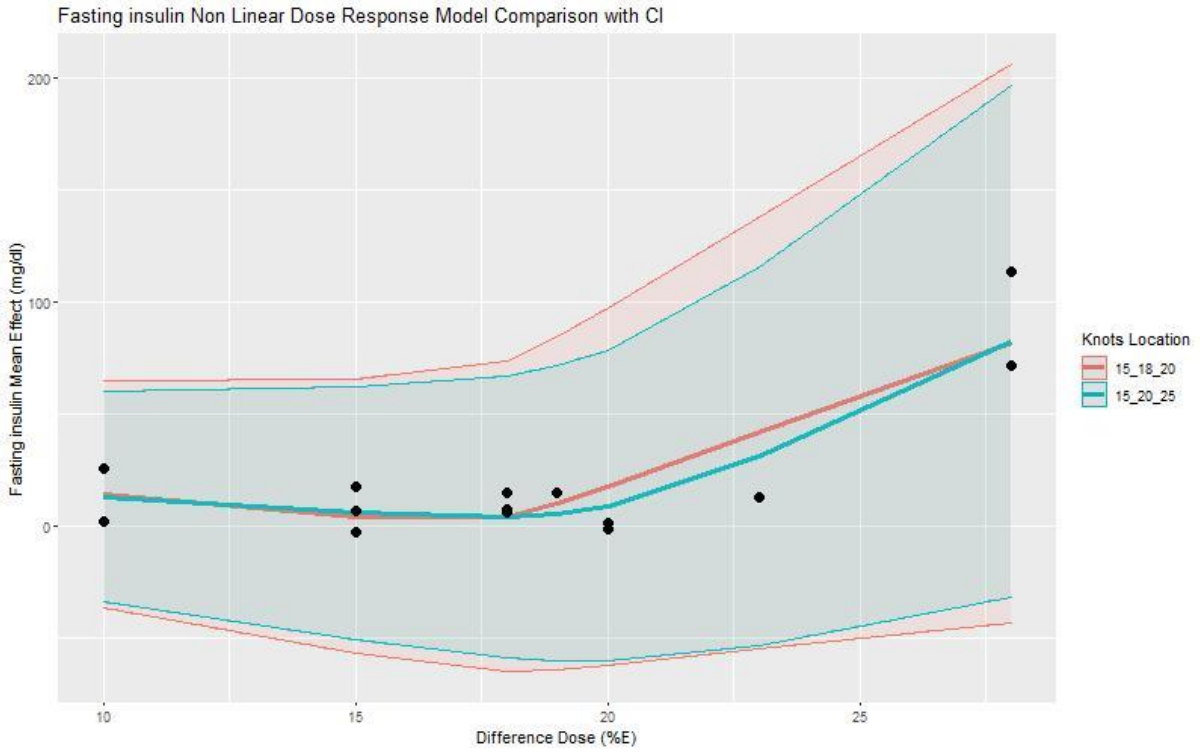
An RCS was used to investigate whether the relationship between the E% sugar intake difference between arms and FI mean effect (i.e. mean difference or mean change difference between arms) could be better described by a non-linear model. In total, 14 observations were eligible for the analysis (after exclusion of the two observations from Moser et al., 1986). The E% sugar intake difference between arms was included as the only fixed effect mimicking the choice done for the linear model. This made the results of the linear and non-linear shape comparable. Two random effects were included to account for the two hierarchical levels in the data (i.e. studies and sex subgroups). The compound symmetry structure was used to describe the correlation between subgroups in the same study.

To investigate the influence of the knot location, the model was run for two of those giving results for AIC, Higgins I^2 and parameter estimates as described in Table 8. The knots location corresponds to different levels of sugar intake difference (E%).

Table 8: Fasting insulin – non-linear model (RCS) estimates at different knots locations

Model	Knots location	AIC	Residual het I^2	Intercept (95% CI)	Beta1 (95% CI)	Beta2 (95% CI)
1	15, 20, 25	102.53	91.9%	26.89 (−1.19, 54.97)	−1.38 (−3.28, 0.51)	7.89 (5.10, 10.67)
2 (centiles 25%, 50%, 75%)	15, 18, 20	102.78	92.6%	34.07 (3.92, 64.22)	−1.98 (−4.05, 0.09)	5.53 (3.57, 7.49)

The two model curves and the related 95% CI of the predicted mean effects on FI are displayed in Figure 18.



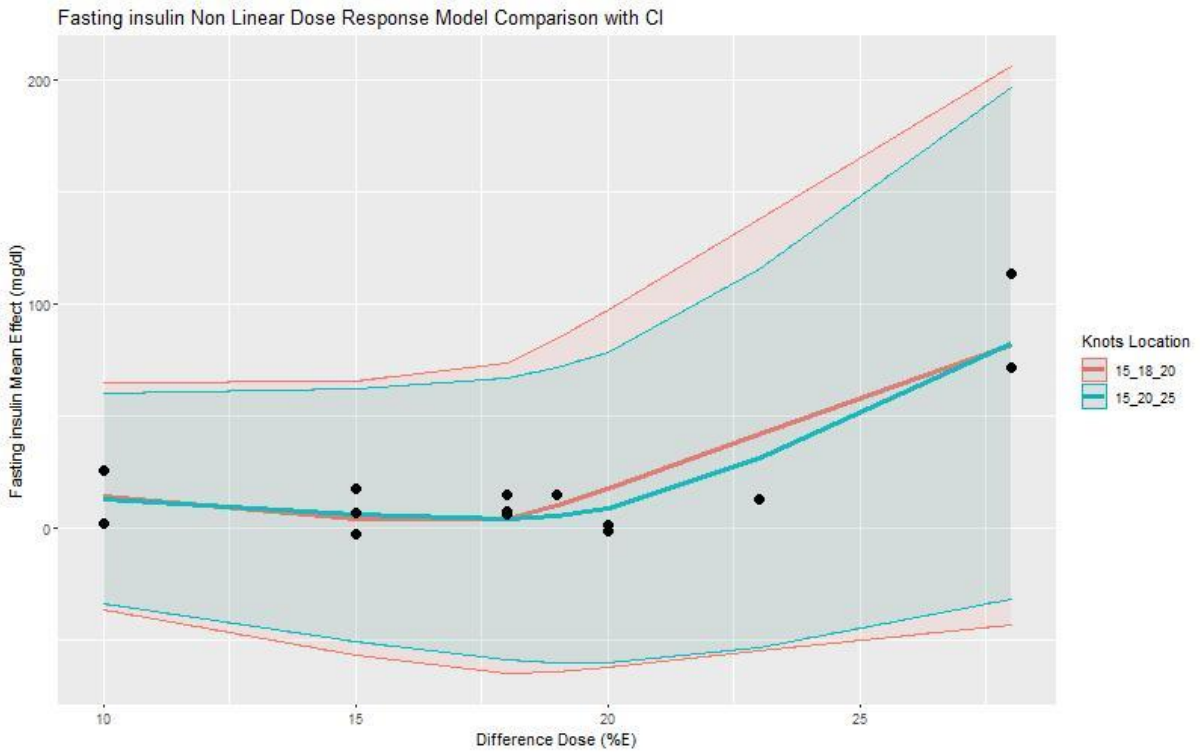


Figure 18: Fasting insulin – comparison of non-linear dose–response models with different knot locations

For the two models the residual heterogeneity remained high (above 90%) and the relative fit was similar (AIC ~102). The 95% CI of the estimate of $\boldsymbol{\beta}_2$ does not overlap zero for any of the models, indicating that a non-linear shape of dose–response relationship might be justified. The estimated non-linear dose–response curve is non-monotonic, slightly decreasing for low E% sugar intakes and increasing with a fast pace at higher E% sugar intake. However, the residual heterogeneity remains high, as well as the sampling uncertainty. In addition, the estimate and precision of the parameter expressing the model non-linear component is highly sensitive to the influence of a study (Israel et al., 1983) (RefID = 5047 in Figure 19). Therefore, the model cannot be used for drawing conclusions, not even on the shape of the dose–response relationship.

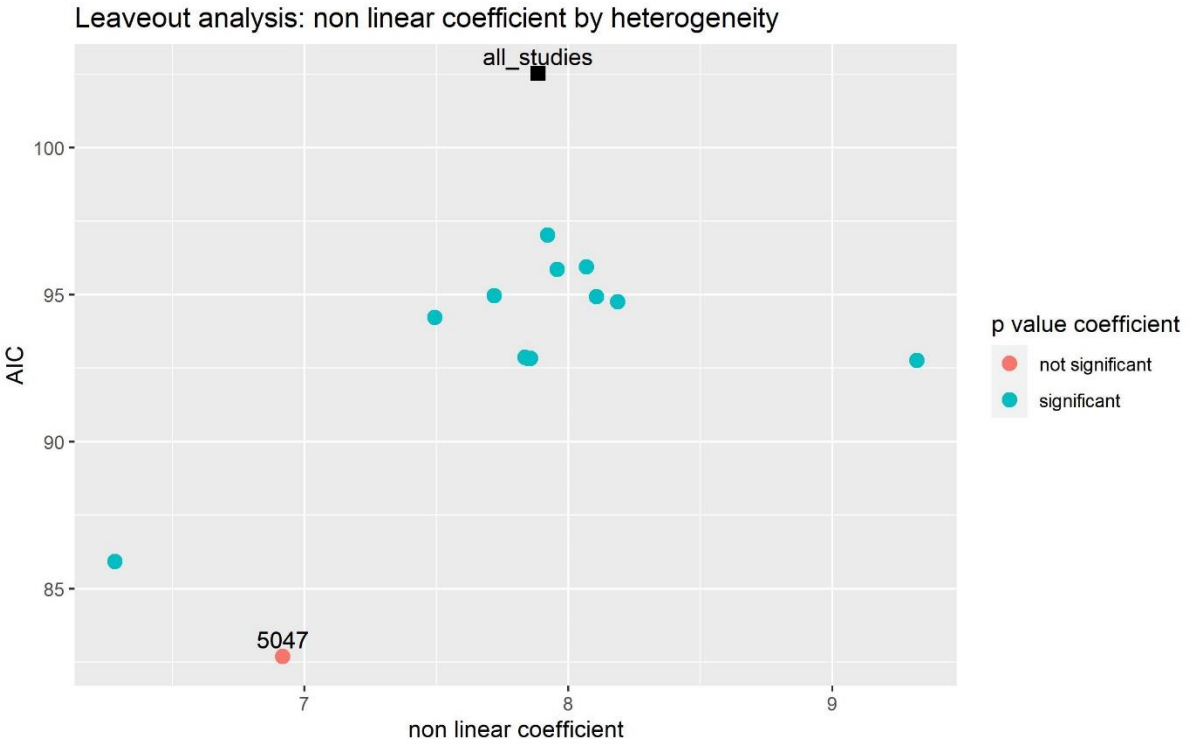


Figure 19: Model diagnostic: leave out analysis for the non-linear component parameter

1.6.5. Meta-regressive non-linear dose–response analysis: body weight (Q1)

An RCS was used to investigate whether the relationship between the E% sugar intake difference between arms and BW mean effect (i.e. mean difference or mean change difference between arms) could be better described by a non-linear model.

In total, 11 observations were eligible for the analysis. The intake of sugars expressed as E% explains a significant component of the variability in the between-arm differences in BW changes (Higgins I^2 for residual heterogeneity = 15.5%). However, the fit of the model measured by the AIC, equal to 34.84, was not dissimilar to that of the linear model (AIC equal to 36.1) and to that of the model with no explanatory variables (AIC equal to 36.5). Moreover, the regression parameter expressing the non-linear effect of the sugar intake on the BW change is extremely low and its CI overlaps zero (0.1491; 95% CI: −0.1522, 0.4504). Due to previous considerations, the impact of other variables as possible modifiers of the effect was not explored. Therefore, evidence does not support a non-linear dose–response relationship between the intake of sugars as E% *ad libitum* and bodyweight.

1.6.6. Meta-regressive non-linear dose–response analysis: uric acid (Q2)

Due to the small number of observations (only five) the non-linear model parameters could not be estimated.

1.6.7. Additional sources of uncertainty not addressed in the models

Two types of uncertainties remained unaddressed. The first one is the compliance to the planned administration. The impact of this source of uncertainty is difficult to predict as it could lead both to an overestimation or underestimation of the E% sugar intake difference. Also it cannot be ignored that compliance differed by treatment groups. Secondly, aggregated data were used to assess the positive

causal relationship between sugar intake and risk of metabolic diseases. It remains uncertain whether similar conclusions would have been reached using individual data.

1.7. Software

Data editing and cleaning was performed using SAS version 9.3. Statistical analyses were carried out in R version 3.3.2 (R Core Team, 2013) and RStudio version 1.0.136. Data cleaning and standardisation were carried out using the 'dplyr' package (Wickham et al., 2021). The meta-regressive linear dose–response analyses were performed using the 'metafor' package (Viechtbauer, 2010), while the non-linear dose–responses were performed using the 'mixmeta' package (Sera et al., 2019). The forest plots were produced using the 'meta' package (Balduzzi et al., 2019). For all the other plots the 'ggplot2' package (Wickham, 2016) was used.

References

Aeberli I, Hochuli M, Gerber PA, Sze L, Murer SB, Tappy L, Spinas GA and Berneis K, 2013. Moderate amounts of fructose consumption impair insulin sensitivity in healthy young men: a randomized controlled trial. Diabetes Care, 36(1), 150–156. doi:10.2337/dc12-0540. Epub August 28 2012.

Akaike H, 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. doi:10.1109/TAC.1974.1100705

Angelopoulos TJ, Lowndes J, Sinnett S and Rippe JM, 2015. Fructose containing sugars do not raise blood pressure or uric acid at normal levels of human consumption. Journal of Clinical Hypertension, 17(2), 87–94. doi:10.1111/jch.12457. Epub December 15 2014.

Balduzzi S, Rücker G and Schwarzer G, 2019. How to perform a meta-analysis with R: a practical tutorial. Evidence-Based Mental Health, 22(4), 153–160. doi:10.1136/ebmental-2019-300117

Bantle JP, Raatz SK, Thomas W and Georgopoulos A, 2000. Effects of dietary fructose on plasma lipids in healthy subjects. American Journal of Clinical Nutrition, 72(5), 1128–1134. doi:10.1093/ajcn/72.5.1128

Belsley DA, Kuh E and Welsch RE, 2005. Regression diagnostics: Identifying influential data and sources of collinearity, John Wiley & Sons.

Black RN, Spence M, McMahon RO, Cuskelly GJ, Ennis CN, McCance DR, Young IS, Bell PM and Hunter SJ, 2006. Effect of eucaloric high- and low-sucrose diets with identical macronutrient profile on insulin resistance and vascular risk: a randomized controlled trial. Diabetes, 55(12), 3566–3572. https://doi.org/10.2337/db06-0220

Borenstein M, Hedges LV, Higgins JPT and Rothstein HR, 2009. Introduction to meta-analysis. Hoboken, John Wiley and Sons.

Campos V, Despland C, Brandejsky V, Kreis R, Schneiter P, Chiolero A, Boesch C and Tappy L, 2015. Sugar- and artificially sweetened beverages and intrahepatic fat: A randomized controlled trial. Obesity, 23(12), 2335–2339. doi:10.1002/oby.21310

Cook RD, 1977. Detection of influential observation in linear regression. Technometrics, 19(1), 15–18. doi:10.1080/00401706.1977.10489493

Crippa A and Orsini N, 2016. Dose-response meta-analysis of differences in means. BMC Medical Research Methodology, 16, 91. doi:10.1186/s12874-016-0189-0

Crippa A, Thomas I and Orsini N, 2018. A pointwise approach to dose-response meta-analysis of aggregated data. International Journal of Statistics in Medical Research, 7(2), 25–32. doi:10.6000/1929-6029.2018.07.02.1

Despland C, Walther B, Kast C, Campos V, Rey V, Stefanoni N and Tappy L, 2017. A randomized-controlled clinical trial of high fructose diets from either Robinia honey or free fructose and glucose in healthy normal weight males. Clinical Nutrition ESPEN, 19, 16–22. doi:10.1016/j.clnesp.2017.01.009

Desquilbet L and Mariotti F, 2010. Dose-response analyses using restricted cubic spline functions in public health research. Statistics in Medicine, 29(9), 1037–1057. doi:10.1002/sim.3841

Ebbeling CB, Feldman HA, Chomitz VR, Antonelli TA, Gortmaker SL, Osganian SK and Ludwig DS, 2012. A randomized trial of sugar-sweetened beverages and adolescent body weight. New England Journal of Medicine, 367(15), 1407–1416. doi:10.1056/NEJMoa1203388. Epub September 21 2012.

EFSA (European Food Safety Authority), 2014. Guidance on expert knowledge elicitation in food and feed safety risk assessment. EFSA Journal 2014;12(6):3734, 55 pp. doi:10.2903/j.efsa.2014.3734

Gostner A, Schäffer V, Theis S, Menzel T, Lührs H, Melcher R, Schauber J, Kudlich T, Dusel G, Dorbath D, Kozianowski G and Scheppach W, 2005. Effects of isomalt consumption on gastrointestinal and metabolic parameters in healthy volunteers. British Journal of Nutrition, 94(4), 575–581. doi:10.1079/bjn20051510

Groen JJ, Balogh M, Yaron E and Cohen AM, 1966. Effect of interchanging bread and sucrose as main source of carbohydrate in a low fat diet on the serum cholesterol levels of healthy volunteer subjects. American Journal of Clinical Nutrition, 19(1), 46–58. doi:10.1093/ajcn/19.1.46

Hallfrisch J, Ellwood KC, Michaelis OEt, Reiser S, O'Dorisio TM and Prather ES, 1983. Effects of dietary fructose on plasma glucose and hormone responses in normal and hyperinsulinemic men. Journal of Nutrition, 113(9), 1819–1826. doi:10.1093/jn/113.9.1819

Hernández-Cordero S, Barquera S, Rodríguez-Ramírez S, Villanueva-Borbolla MA, González de Cossio T, Dommarco JR and Popkin B, 2014. Substituting water for sugar-sweetened beverages reduces circulating triglycerides and the prevalence of metabolic syndrome in obese but not in overweight Mexican women in a randomized controlled trial. Journal of Nutrition, 144(11), 1742–1752. doi:10.3945/jn.114.193490. Epub September 3 2014.

Higgins JP and Thompson SG, 2002. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. doi:10.1002/sim.1186

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ and Welch VA (eds.), 2020. Cochrane Handbook for Systematic Reviews of Interventions version 6.1. Cochrane. Available online: www.training.cochrane.org/handbook.

Hollis JH, Houchins JA, Blumberg JB and Mattes RD, 2009. Effects of concord grape juice on appetite, diet, body weight, lipid profile, and antioxidant status of adults. Journal of the American College of Nutrition, 28(5), 574–582. doi:10.1080/07315724.2009.10719789

Huttunen JK, Makinen KK and Scheinin A, 1976. Turku sugar studies XI. Effects of sucrose, fructose and xylitol diets on glucose, lipid and urate metabolism. Acta Odontologica Scandinavica, 34(6), 345–351. doi:10.3109/00016357609004646

Israel KD, Michaelis OEt, Reiser S and Keeney M, 1983. Serum uric acid, inorganic phosphorus, and glutamic-oxalacetic transaminase and blood pressure in carbohydrate-sensitive adults consuming three different levels of sucrose. Annals of Nutrition and Metabolism, 27(5), 425–435. doi:10.1159/000176714

Jin R, Welsh JA, Le NA, Holzberg J, Sharma P, Martin DR and Vos MB, 2014. Dietary fructose reduction improves markers of cardiovascular disease risk in Hispanic-American adolescents with NAFLD. Nutrients, 6(8), 3187–3201. doi:10.3390/nu6083187

Johnston RD, Stephenson MC, Crossland H, Cordon SM, Palcidi E, Cox EF, Taylor MA, Aithal GP and Macdonald IA, 2013. No difference between high-fructose and high-glucose diets on liver triacylglycerol or biochemistry in healthy overweight men. Gastroenterology, 145(5), 1016–1025.e2. doi:10.1053/j.gastro.2013.07.012

Koh ET, Ard NF and Mendoza F, 1988. Effects of fructose feeding on blood parameters and blood pressure in impaired glucose-tolerant subjects. Journal of the American Dietetic Association, 88(8), 932–938. doi:10.1016/S0002-8223(21)07931-1

Lewis AS, McCourt HJ, Ennis CN, Bell PM, Courtney CH, McKinley MC, Young IS and Hunter SJ, 2013. Comparison of 5% versus 15% sucrose intakes as part of a eucaloric diet in overweight and obese subjects: effects on insulin sensitivity, glucose metabolism, vascular compliance, body composition and lipid profile. A randomised controlled trial. Metabolism: Clinical and Experimental, 62(5), 694–702. doi:10.1016/j.metabol.2012.11.008. Epub January 28 2013.

Light, R.J., & Pillemer, D.B. (1984). Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press. xiii+ 191 pp.

Lowndes J, Sinnett S, Pardo S, Nguyen VT, Melanson KJ, Yu Z, Lowther BE and Rippe JM, 2014a. The effect of normally consumed amounts of sucrose or high fructose corn syrup on lipid profiles, body composition and related parameters in overweight/obese subjects. Nutrients, 6(3), 1128–1144. doi:10.3390/nu6031128

Lowndes J, Sinnett S, Yu Z and Rippe J, 2014b. The effects of fructose-containing sugars on weight, body composition and cardiometabolic risk factors when consumed at up to the 90th percentile population consumption level for fructose. Nutrients, 6(8), 3153–3168. doi:10.3390/nu6083153

Lowndes J, Sinnett SS and Rippe JM, 2015. No effect of added sugar consumed at median American intake level on glucose tolerance or insulin resistance. Nutrients, 7(10), 8830–8845. doi:10.3390/nu7105430

Maersk M, Belza A, Stodkilde-Jorgensen H, Ringgaard S, Chabanova E, Thomsen H, Pedersen SB, Astrup A and Richelsen B, 2012. Sucrose-sweetened beverages increase fat storage in the liver, muscle, and visceral fat depot: a 6-mo randomized intervention study. American Journal of Clinical Nutrition, 95(2), 283–289. doi:10.3945/ajcn.111.022533. Epub December 28 2011.

Majid M, Younis MA, Naveed AK, Shah MU, Azeem Z and Tirmizi SH, 2013. Effects of natural honey on blood glucose and lipid profile in young healthy Pakistani males. Journal of Ayub Medical College, Abbottabad, 25(3–4), 44–47.

Marinho VC, Higgins JP, Sheiham A and Logan S, 2003. Fluoride toothpastes for preventing dental caries in children and adolescents. Cochrane Database of Systematic Reviews, (1), CD002278. doi:10.1002/14651858.CD002278

Mark AB, Poulsen MW, Andersen S, Andersen JM, Bak MJ, Ritz C, Holst JJ, Nielsen J, de Courten B, Dragsted LO and Bugel SG, 2014. Consumption of a diet low in advanced glycation end products for 4 weeks improves insulin sensitivity in overweight women. Diabetes Care, 37(1), 88–95. doi:10.2337/dc13-0842. Epub August 19 2013.

Markey O, Le Jeune J and Lovegrove JA, 2016. Energy compensation following consumption of sugar-reduced products: a randomized controlled trial. European Journal of Nutrition, 55(6), 2137–2149. doi:10.1007/s00394-015-1028-5. Epub September 9 2015.

Moser PB, Behall KM, Kelsay JL and Prather ES, 1986. Carbohydrate tolerance and serum lipid responses to type of dietary carbohydrate and oral contraceptive use in young women. Journal of the American College of Nutrition, 5(1), 45–53. doi:10.1080/07315724.1986.10720111

R Core Team, 2013. R: A language and environment for statistical computing. Available online: http://www.r-project.org/

Raben A, Vasilaras TH, Christina Møller A and Astrup A, 2002. Sucrose compared with artificial sweeteners: Different effects on *ad libitum* food intake and body weight after 10 wk of supplementation in overweight subjects. American Journal of Clinical Nutrition, 76(4), 721–729. doi:10.1093/ajcn/76.4.721

Reiser S, Hallfrisch J, Michaelis OE, Lazar FL, Martin RE and Prather ES, 1979. Isocaloric exchange of dietary starch and sucrose in humans. I. Effects on levels of fasting blood lipids. American Journal of Clinical Nutrition, 32(8), 1659–1669. doi:10.1093/ajcn/32.8.1659

Reiser S, Powell AS, Scholfield DJ, Panda P, Ellwood KC and Canary JJ, 1989. Blood lipids, lipoproteins, apoproteins, and uric acid in men fed diets containing fructose or high-amylose cornstarch. American Journal of Clinical Nutrition, 49(5), 832–839. doi:10.1093/ajcn/49.5.832

Ruyter de JC, Olthof MR, Seidell JC and Katan MB, 2014. A trial of sugar-free or sugar-sweetened beverages and body weight in children. World Review of Nutrition and Dietetics, 109, 4–5. doi:10.1159/000356352

Saltelli A, Chan K and Scott EM, 2009. Sensitivity analysis. New York, John Wiley & Sons.

Saris WH, Astrup A, Prentice AM, Zunft HJ, Formiguera X, Verboeket-van de Venne WP, Raben A, Poppitt SD, Seppelt B, Johnston S, Vasilaras TH and Keogh GF, 2000. Randomized controlled trial of changes in dietary carbohydrate/fat ratio and simple vs complex carbohydrates on body weight and blood lipids: the CARMEN study. The Carbohydrate Ratio Management in European National diets. International Journal of Obesity and Related Metabolic Disorders, 24(10), 1310–1318. doi:10.1038/sj.ijo.0801451

Schwarz JM, Noworolski SM, Wen MJ, Dyachenko A, Prior JL, Weinberg ME, Herraiz LA, Tai VW, Bergeron N, Bersot TP, Rao MN, Schambelan M and Mulligan K, 2015. Effect of a high-fructose weight-maintaining diet on lipogenesis and liver fat. Journal of Clinical Endocrinology and Metabolism, 100(6), 2434–2442. doi:10.1210/jc.2014-3678

Sera F, Armstrong B, Blangiardo M and Gasparrini A, 2019. An extended mixed-effects framework for meta-analysis. Statistics in Medicine, 38(29), 5429–5444. doi:10.1002/sim.8362

Shapiro SS and Wilk MB, 1965. An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. doi:10.1093/biomet/52.3-4.591

Smith JB, Niven BE and Mann JI, 1996. The effect of reduced extrinsic sucrose intake on plasma triglyceride levels. European Journal of Clinical Nutrition, 50(8), 498–504.

Stanhope KL, Schwarz JM, Keim NL, Griffen SC, Bremer AA, Graham JL, Hatcher B, Cox CL, Dyachenko A, Zhang W, McGahan JP, Seibert A, Krauss RM, Chiu S, Schaefer EJ, Ai M, Otokozawa S, Nakajima K, Nakano T, Beysen C, Hellerstein MK, Berglund L and Havel PJ, 2009. Consuming fructose-sweetened, not glucose-sweetened, beverages increases visceral adiposity and lipids and decreases insulin sensitivity in overweight/obese humans. Journal of Clinical Investigation, 119(5), 1322–1334. doi:10.1172/JCI37385

Surowska A, Jegatheesan P, Campos V, Marques A-S, Egli L, Cros J, Rosset R, Lecoultre V, Kreis R, Boesch C, Pouymayou B, Schneiter P and Tappy L, 2019. Effects of dietary protein and fat content on intrahepatocellular and intramyocellular lipids during a 6-day hypercaloric, high sucrose diet: a randomized controlled trial in normal weight healthy subjects. Nutrients, 11(1), 209. doi:10.3390/nu11010209

Swanson JE, Laine DC, Thomas W and Bantle JP, 1992. Metabolic effects of dietary fructose in healthy subjects. American Journal of Clinical Nutrition, 55(4), 851–856. doi:10.1093/ajcn/55.4.851

Szanto S and Yudkin J, 1969. The effect of dietary sucrose on blood lipids, serum insulin, platelet adhesiveness and body weight in human volunteers. Postgraduate Medical Journal, 45(527), 602–607. doi:10.1136/pgmj.45.527.602

Thompson RG, Hayford JT and Danney MM, 1978. Glucose and insulin responses to diet. Effect of variations in source and amount of carbohydrate. Diabetes, 27(10), 1020–1026. doi:10.2337/diab.27.10.1020

Umpleby AM, Shojaee-Moradie F, Fielding B, Li X, Marino A, Alsini N, Isherwood C, Jackson N, Ahmad A, Stolinski M, Lovegrove JA, Johnsen S, Jeewaka RMAS, Wright J, Wilinska ME, Hovorka R, Bell JD, Thomas EL, Frost GS and Griffin BA, 2017. Impact of liver fat on the differential partitioning of hepatic triacylglycerol into VLDL subclasses on high and low sugar diets. Clinical Science (London), 131(21), 2561–2573. doi:10.1042/CS20171208

Valtueña S, Pellegrini N, Franzini L, Bianchi MA, Ardigo D, Del Rio D, Piatti P, Scazzina F, Zavaroni I and Brighenti F, 2008. Food selection based on total antioxidant capacity can modify antioxidant intake, systemic inflammation, and liver function without altering markers of oxidative stress. American Journal of Clinical Nutrition, 87(5), 1290–1297. doi:10.1093/ajcn/87.5.1290

Viechtbauer W, 2010. metafor: Meta-analysis package for R. R package version 1.4-0. URL https://CRAN.R-project.org/package=metafor

Werner D, Emmett PM and Heaton KW, 1984. Effects of dietary sucrose on factors influencing cholesterol gall stone formation. Gut, 25(3), 269–274. doi:10.1136/gut.25.3.269

Wickham H, 2016. ggplot2: Elegant graphics for data analysis, Springer-Verlag, New York.

Wickham H, François R, Henry L and Müller K, 2021. dplyr: Adplyr: Grammar of Data Manipulation. Retrieved from R package version 1.0.4. Available online: https://CRAN.R-project.org/package=dplyr

Wilk MB and Gnanadesikan R, 1968. Probability plotting methods for the analysis for the analysis of data. Biometrika, 55(1), 1–17. doi:10.1093/biomet/55.1.1

Appendix A – Measurement units conversion factors

Endpoint	Old unit	Conversion factor	Code unit	Final unit
Body weight	Kg	1	10	kg
	lb	0.45	11	kg
Body fat	%	1	9	%
BMI	kg/m^2	1	12	kg/m^2
Waist circumference	cm	1	13	cm
Clamp – hepatic	mg/kg/min	1	2	mg/kg/min
	µmol/kg/min	21176.47	3	mg/kg/min
Clamp – whole body	mL/kg/min		1	
	µmol/kg/min	21176.47	2	mg/kg/min
	mg/kg/min	1	3	mg/kg/min
	mg/kg/min/µU/dL		4	
Clamp – insulin	mU /kg/min		1	
	$mU/m^2/min$		2	
Total cholesterol	mmol/L	38.67	3	mg/dL
	mg/dL	1	1	mg/dL
	mg/100 ml	1	2	mg/dL
	mg/L	0.1	5	mg/dL
HDL cholesterol	mmol/L	38.67	3	mg/dL
	mg/dL	1	1	mg/dL
	mg/100 ml	1	2	mg/dL
	mg/L	0.1	5	mg/dL
LDL cholesterol	mmol/L	38.67	3	mg/dL
	mg/dL	1	1	mg/dL
	mg/100 ml	1	2	mg/dL
	mg/L	0.1	5	mg/dL
Ectopic fat VAT	kg		10	
	cm^3 (e.g. VAT)		14	
Ectopic fat – liver fat	mmol/L		3	
	%		9	
	% from baseline		15	
	AU (arbitrary units)		16	
	% signal		17	
IVITT	10^{-2} min^{-1}	1	15	10^{-2} min^{-1}
OGTT – glucose	mg/100 ml	1	1	mg/dL
	mg/dL	1	2	mg/dL
	mmol/L	18	3	mg/dL
OGTT – insulin	µU/mL	6.94	1	pmol/L
	mU/L	6.94	3	pmol/L
	pmol/L	1	5	pmol/L
	µg/L	200.144	7	pmol/L

Fasting glucose	mg/dL	1	1	mg/dL
	mg/100 ml	1	2	mg/dL
	mmol/L	18	3	mg/dL
Fasting insulin	mmol/L	1.00E+09	3	pmol/L
	µU/mL	6.94	4	pmol/L
	µU/mL	6.94	5	pmol/L
	mU/L	6.94	6	pmol/L
	pmol/L	1	16	pmol/L
	µg/L	200.144	15	pmol/L
Triglycerides	mg/dL	1	1	mg/dL
	mg/100 mL	1	2	mg/dL
	mmol/L	88.5	3	mg/dL
	mg/L	0.1	15	mg/dL
Uric acid	mg/dL	1	1	mg/dL
	mg/100 mL	1	2	mg/dL
	mmol/L	16.81	3	mg/dL
	µmol/L	0.01681	8	mg/dL
	mg/L	0.1	15	mg/dL

Appendix B – Control–intervention arm identification by study

Table B1: Control and intervention arm classification for each respective question Q1 and Q2 (c = control arm; i = intervention arm; x = not relevant for the question)

Author	Arms	Sugars dose [E%]	Source	Q1	Q2
Aeberli et al. (2013)	Fructose	8	Beverages	c	x
Aeberli et al. (2013)	Fructose	16	Beverages	i	i
Aeberli et al. (2013)	Glucose	16	Beverages	x	c
Aeberli et al. (2013)	Sucrose	16	Beverages	x	x
Angelopoulos et al. (2015)*	HFCS	18	Beverages	x	x
Angelopoulos et al. (2015)*	Sucrose	18	Beverages	x	x
Angelopoulos et al. (2015)*	Fructose	9	Beverages	x	i
Angelopoulos et al. (2015)*	Glucose	9	Beverages	x	c
Bantle et al. (2000)	Fructose	14	Mixed	x	i
Bantle et al. (2000)	Glucose	14	Mixed	x	c
Black et al. (2006)	Sucrose	25	Mixed	i	x
Black et al. (2006)	Sucrose	10	Mixed	c	x
Campos et al. (2015)	ASSD	0	Beverages	c	x
Campos et al. (2015)	SSSD	18	Beverages	i	x
Despland et al. (2017)	Starch	0	Mixed	c	x
Despland et al. (2017)	Honey	25	Mixed	x	x
Despland et al. (2017)	Glucose/fructose	25	Mixed	i	x
Ebbeling et al. (2012)	SSSD+SSFD+TFJ	17	Beverages	i	x
Ebbeling et al. (2012)	ASSD+water	0	Beverages	c	x
Gostner et al. (2005)	Isomalt	0	Foods	c	x
Gostner et al. (2005)	Sucrose	6	Foods	i	x
Groen et al. (1966)	Starch	0	Mixed	c	x
Groen et al. (1966)	Sucrose	30	Mixed	i	x
Hallfrisch et al. (1983)a*	Starch	0	Foods	c	x
Hallfrisch et al. (1983)a*	Fructose	7.5	Foods	x	x
Hallfrisch et al. (1983)a*	Fructose	15	Foods	i	x
Hernández-Cordero et al. (2014)	Water	0	Beverages	c	x
Hernández-Cordero et al. (2014)	SSBs	20	Beverages	i	x
Hollis et al. (2009)	Grape juice	18	Beverages	x	x
Hollis et al. (2009)	Grape drink	18	Beverages	i	x
Hollis et al. (2009)	No beverage	0	Beverages	c	x
Huttunen et al. (1976)	Xylitol	0	Mixed	c	x
Huttunen et al. (1976)	Sucrose	16	Mixed	i	x
Huttunen et al. (1976)	Fructose	14	Mixed	x	x
Israel et al. (1983)*	Sucrose	2	Foods	c	x
Israel et al. (1983)*	Sucrose	15	Foods	x	x
Israel et al. (1983)*	Sucrose	30	Foods	i	x

Jin et al. (2014)	Fructose	20	Beverages	x	i
Jin et al. (2014)	Glucose	20	Beverages	x	c
Johnston et al. (2013)	Fructose	25	Beverages	x	i
Johnston et al. (2013)	Glucose	25	Beverages	x	c
Koh et al. (1988)	Fructose	15	Mixed	x	i
Koh et al. (1988)	Glucose	15	Mixed	x	c
Lewis et al. (2013)	Sucrose	5	Mixed	c	x
Lewis et al. (2013)	Sucrose	15	Mixed	i	x
Lowndes et al. (2014a)	HFCS	10	Beverages	x	x
Lowndes et al. (2014a)	HFCS	20	Beverages	x	x
Lowndes et al. (2014a)	Sucrose	10	Beverages	c	x
Lowndes et al. (2014a)	Sucrose	20	Beverages	i	x
Lowndes et al. (2014b)*	Sucrose	8	Beverages	c	x
Lowndes et al. (2014b)*	HFCS	8	Beverages	x	x
Lowndes et al. (2014b)*	Sucrose	18	Beverages	x	x
Lowndes et al. (2014b)*	HFCS	18	Beverages	x	x
Lowndes et al. (2014b)*	HFCS	30	Beverages	x	x
Lowndes et al. (2014b)*	Sucrose	30	Beverages	i	x
Lowndes et al. (2015)	HFCS	18	Beverages	x	x
Lowndes et al. (2015)	Fructose	9	Beverages	x	i
Lowndes et al. (2015)	Glucose	9	Beverages	x	c
Lowndes et al. (2015)	Sucrose	18	Beverages	i	x
Lowndes et al. (2015)	Control milk	0	Beverages	c	x
Maersk et al. (2012)*	SSSD	18	Beverages	i	x
Maersk et al. (2012)*	Water	0	Beverages	x	x
Maersk et al. (2012)*	ASSD	0	Beverages	c	x
Majid et al. (2013)	Honey	8	Beverages	i	x
Majid et al. (2013)	No beverage	0	Beverages	c	x
Mark et al. (2014)	Fructose	14	Beverages	x	i
Mark et al. (2014)	Glucose	16	Beverages	x	c
Markey et al. (2016)	NMES	16	Mixed	i	x
Markey et al. (2016)	NMES	6	Mixed	c	x
Moser et al. (1986)	Sucrose	43	Foods	i	x
Moser et al. (1986)	Starch	0	Foods	c	x
Raben et al. (2002)*	Sucrose	23	Mixed	i	x
Raben et al. (2002)*	Artificial sweeteners	0	Mixed	c	x
Reiser et al. (1979)*	Sucrose	30	Foods	i	x
Reiser et al. (1979)*	Starch	0	Foods	c	x
Reiser et al. (1989)*	Sucrose	20	Foods	i	x
Reiser et al. (1989)*	Starch	0	Foods	c	x
Ruyter de et al. (2014)	SSSD	6	Beverages	i	x
Ruyter de et al. (2014)	ASSD	0	Beverages	c	x
Saris et al. (2000)*	High simple CHO	38	Mixed	i	x

Saris et al. (2000)*	High complex CHO	19	Mixed	c	x
Saris et al. (2000)*	Control	22.3	Mixed	x	x
Schwarz et al. (2015)	Starch	0	Beverages	c	x
Schwarz et al. (2015)	Fructose	20	Beverages	i	x
Smith et al. (1996)	Sucrose	12	Mixed	i	x
Smith et al. (1996)	Sugar-free diet	0	Mixed	c	x
Stanhope et al. (2009)*	Glucose	25	Beverages	x	c
Stanhope et al. (2009)*	Fructose	25	Beverages	x	i
Swanson et al. (1992)	Fructose	16.6	Mixed	i	x
Swanson et al. (1992)	Fructose	0	Mixed	c	x
Szanto and Yudkin (1969)	Sucrose	54	Mixed	i	x
Szanto and Yudkin (1969)	Starch	0	Mixed	c	x
Thompson et al. (1978)	Sucrose	45	Beverages	c	x
Thompson et al. (1978)	Sucrose	65	Beverages	i	x
Thompson et al. (1978)	Corn syrup	45	Beverages	x	x
Thompson et al. (1978)	Corn syrup	65	Beverages	x	x
Umpleby et al. (2017)	NMES	26	Mixed	i	x
Umpleby et al. (2017)	NMES	6	Mixed	c	x
Werner et al. (1984)	Sucrose	24	Mixed	i	x
Werner et al. (1984)	Artificial sweeteners	0	Mixed	c	x

ASSD: artificially sweetened soft drinks; CHO: carbohydrates; HFCS: high-fructose corn syrup; NMES: non-mil extrinsic sugars; SSB: sugar-sweetened beverages; SSFD: sugar-sweetened fruit drinks, SSSD: sugar-sweetened soft drinks; TFJ: total fruit juice. * Data for some endpoints extracted from linked reference

Appendix C – Report of EKE for correlation coefficient

C.1. Background

The mean effects measure (i.e. the difference between the achieved mean level or the change from baseline of the endpoint in the intervention and the control arm) and its precision are generally not reported in studies and had to be to be computed. Generally, the study authors reported the endpoint mean at baseline (mainly for parallel studies) and at the end of the treatment by treatment groups and the relationship to SD or SE. The method used to compute the mean effects in parallel and cross-over studies is described in Section 1.1.2. The computation of the precision of the mean effect entailed the estimation of a correlation coefficient for the reasons described in Section 1.1.4.

The correlation coefficients between baseline and post-treatment measurements within groups varied considerably across endpoints and between study groups, in studies for which such correlations could be calculated. Therefore there is a large uncertainty on what the true value of the correlation coefficient between baseline and post-treatment would be in different study conditions and for different endpoints.

C.2. Effect measure: precision of the estimate (SE)

For parallel studies, when the achieved levels at the end of the treatment is compared, the SE of the mean effect can be easily derived from the SD or SE under the assumption of independence of the observations across groups. In principle, the latter assumption is not completely met since individuals are not randomly sampled from the population and measurements were taken under the same experimental setting. However, for sake of simplicity and considering that most probably the correlation would be very low, the assumption is considered acceptable. There is no need to estimate a correlation coefficient across different treatment groups. This is necessary, however, when the change from baseline is compared across groups, as the endpoints are measured on the same individuals at the beginning and at the end of the trial.

Similarly for cross-over studies, the assumption of independence across groups does not hold, as the two treatments are administered to the same individuals at different times. Depending on the endpoint, the treatment and its duration, the bias applied to the estimate of the SD/SE assumes that independence might be very large. Generally, study authors reported the SD/SE of the endpoint post-treatment means in each group. To estimate the SD/SE of the effect measure, an estimate of the correlation coefficient is needed.

C.3. Direction/level of conservativism

The impact of ignoring the correlation across groups in a cross-over study would be a bias in the estimate of the SE (precision) or the estimate of the effect measure. This bias would imply an underestimation of the precision (overestimation of the CI width) in case the true correlation is positive and greater than 0.50. A similar impact would have an underestimation of a positive correlation as shown in Figures 1 and 2.

The opposite impact would be observed if the correlation was negative.

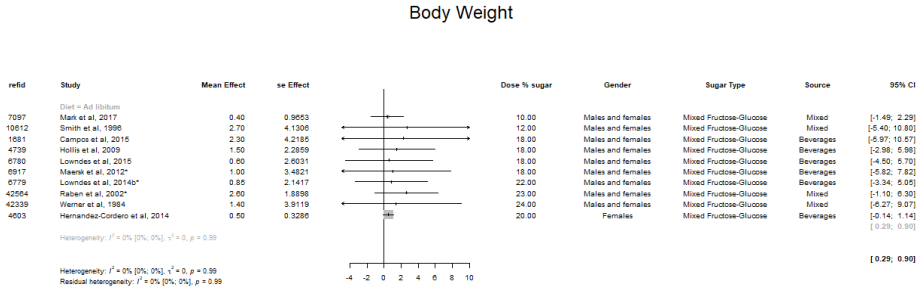
We will focus the following discussion on the assumption that a positive correlation exists, as supported by the limited evidence collected for this assessment on the issue.

In case of a positive correlation, the consequence of underestimating/overestimating the correlation and then the precision would be to:

- increase/decrease the probability NOT TO SEE a significant effect of the treatment on the endpoint when the effect IS TRUE (FALSE NEGATIVE) → beta error;
- decrease/increase the probability TO SEE a significant effect of the treatment on the endpoint when the effect IS NOT TRUE (FALSE POSITIVE) → alpha error.

The WG was requested to decide on what would be the most serious bias/error (FALSE POSTIVE or FALSE NEGATIVE). The conclusion was that a false negative was the most serious mistake.

Correlation coefficient= 0.70



Correlation coefficient = 0.90

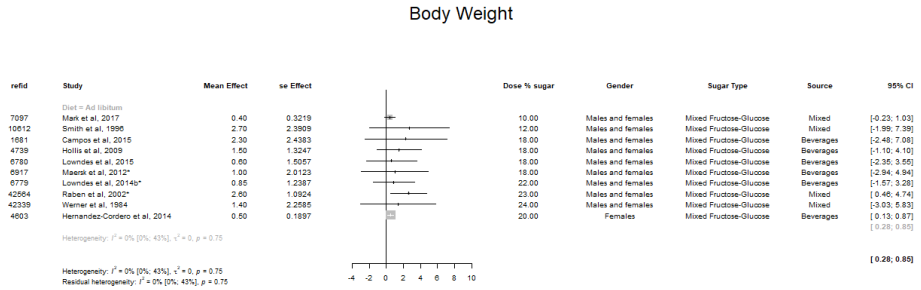


Figure C1: Examples of how the correlation coefficient affects the precision (95% CI) of the mean effect estimate (and therefore potentially the significance of the effect), but not the mean estimate

C.4. Level of accuracy in the estimate of the correlation coefficient

The true value of the correlation of the post-treatment endpoint mean in a cross-over trial at the end of each treatment can vary significantly in the real world depending on:

- type of intervention
- duration of the treatment and the wash-out
- population and its sensitivity to the treatment
- endpoint
- setting.

Although it would be desirable to have an estimate of the correlation for each type of endpoint and even for each single study, this approach would be time consuming with estimates still prone to a large uncertainty, considering the limited evidence at hand.

C.5. Approach used for the elicitation of the correlation coefficient(s)

Given the high level of uncertainty in the level of correlation in cross-over studies and the limited evidence that is available to provide an accurate and precise estimate for our BoE, the following approach is proposed:

- Identify a single range that would cover with 95% probability the true level of the correlation coefficient for all the metabolic endpoints observed after intake of different levels of sugar in a population similar to our target population.

- Identify within the range a 'best estimate' for the correlation coefficient in cross-over studies.

On average, some endpoints will be more stable than others and less affected by treatment, its duration, etc. (e.g. body weight vs. biochemistry variables such as FG or fasting triglycerides). Consequently, on average the correlation coefficient between post-treatment levels is expected to be higher for these endpoints.

The best estimate will be used for the computation of the SE of the effect measure for displaying the confidence intervals in the forest plots and for the computation of the dose–response meta-analysis on fasting triglycerides, BW and FG.

In all cases, a sensitivity analysis will be undertaken, using the upper and lower boundaries of the range, to assess the influence of the uncertainty in the correlation coefficient estimate on the final results.

Considering the limited time available at the meeting, a credible range for the correlation coefficient that would cover with 95% probability the true level of the parameter for all the metabolic endpoints observed after intake of different levels of sugar in a population similar to our target population was estimated by the staff and proposed to the WG for confirmation. The initial proposed range goes from 0.60 to 0.99. It was derived on the basis of two studies Valtueña et al. (2008) and Surowska et al. (2019) whose individual data were provided by the authors for the exclusive scope of the EKE.

C.6. Pre-meeting questions sent to the experts before the meeting to agree on preliminary aspects

Before the elicitation meeting, the experts were required by email to answer individually to the questions listed in Table 9.

Table 9: Questions answered by the experts individually

Questions	Answer a)	Answer b)
What do you think would be a more serious bias/error in the context of this safety assessment	NOT TO SEE an effect when the effect IS TRUE (FALSE NEGATIVE)	TO SEE and effect when the effect IS NOT TRUE (FALSE POSITIVE)
What do you think about having a single value for the correlation coefficient for all cross-over studies and endpoint variables:	Agree	Not agree. If so please provide an alternative with a reason
Do you think that the range having a 95% probability of including the TRUE value for the correlation coefficient across variables and studies could be between 0.60 and 0.99?	Agree	Not agree. If so, please provide a new range with a reason, e.g. other evidence or data you might have, so this can be shared with others

C.7. Pre-meeting expert knowledge elicitation question and evidence dossier

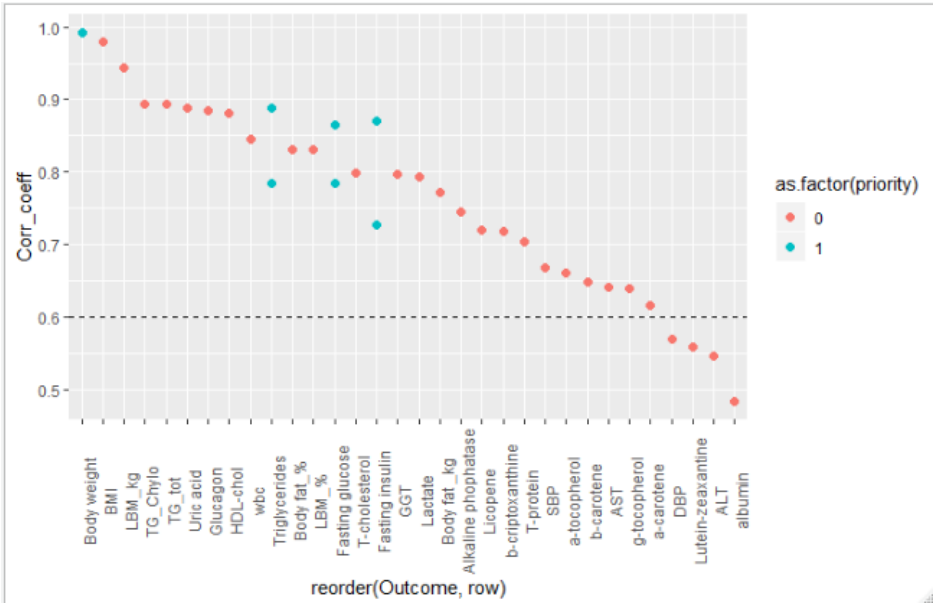
The quantity to be estimated and the related questions are provided in Table 10.

Table 10: Quantity to be estimated and related questions

Estimated quantity	Question

Estimate of intra-individual correlation coefficient in cross-over trials (post-treatment 1 vs post-treatment 2)	Think of the population of cross-over trials in which two different levels of sugars are administered in sequence to a large sample of individuals representative of a healthy population. Assume that a wash-out period of sufficient duration is applied after the first treatment. Also assume that the treatment is not leading to any carry-over. Doses can be administered in any order. 'Correlation coefficient' means the parameter that would describe how similar would be, on average in a population of cross-over trials (similar to the ones considered in our assessment), the post-treatment measurements of each endpoint taken after administering in sequence two different doses of sugars to a large sample of individuals representative of a healthy population. As agreed, it is assumed that a credible range for the correlation coefficient (as defined above) that would cover with 95% probability the true average value of the correlation, in the population of cross-over studies, is between 0.60 and 0.99 Provide your best estimate of the correlation coefficient considering the following: • The evidence at hand (in annex A) is extremely limited and is purely indicative. • The type of error that you considered more serious is not detecting an effect when it is true (FALSE NEGATIVE). Consequently an overestimation of the parameter would be less serious of an underestimation since it would decrease probability of false negative. • Focus on the endpoints that will be analysed quantitatively: body weight, fasting triglycerides, fasting glucose and insulin but think of a value that would cover also the other endpoints (in case with a greater bias). • Bear in mind that the best estimate is aimed at minimising on average the bias in the estimate of the correlation coefficient for the three most relevant endpoints (body weight, fasting triglycerides, fasting glucose and insulin).

C.8. Material provided at the meeting (13 December 2019)



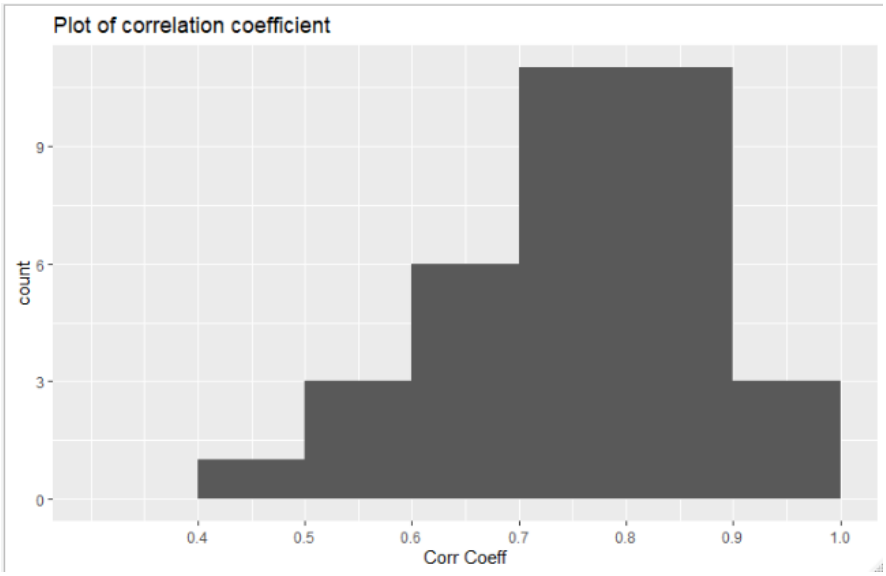


Figure C2: Summary of the correlation coefficients estimated in the two available datasets

Quantiles of the distribution of correlation coefficient values for endpoints 1 (fasting triglycerides, fasting glucose and insulin, body weight)

5%	25%	50%	Mean	75%	95%
0.7444	0.7850	0.8660	0.85	0.8790	0.9608

Quantiles of the distribution of correlation coefficient values for endpoints 2 (all but endpoint 1)

5%	25%	50%	Mean	75%	95%
0.55185	0.64700	0.75900	0.75	0.85425	0.92585

C.9. Expert knowledge elicitation process

Before the start of the elicitation process, the following aspects were discussed:

- Whether all experts could agree that FALSE NEGATIVE is the most serious error in our context. One expert was not in line with the others but could easily agree at the end.
- Whether the range had to be extended up to a lower boundary of 0.50, as suggested by one expert. The rest of the group agreed on this proposal listening to the motives.

Therefore the original range for the correlation coefficient was revised and fixed at (0.50–0.99).

A brief introduction was given on the criteria to use to provide the estimate (criteria above) and a clarification was provided on the impact of the over- or underestimation on the probability of a false negative. The larger the underestimation, the larger the probability of a false negative.

Then the experts provided their 'best estimate'.

Expert	Best estimate
1	0.75
2	0.85
3	0.88
4	0.76
5	0.87
Average	0.82

Since the estimates were similar (range between 0.75 and 0.88) it was agreed to take the arithmetic average as the consensus estimate.

Therefore the best estimate for the correlation coefficient was fixed at 0.82.

It was agreed also to use as the uncertainty distribution for the correlation coefficient in cross-over trials the Pert distribution with parameters as provided by the experts (min. = 0.5, max. = 0.99, mode = 0.82).

5%	25%	50%	Mean	75%	95%
0.50	0.73	0.80	0.80	0.86	0.99

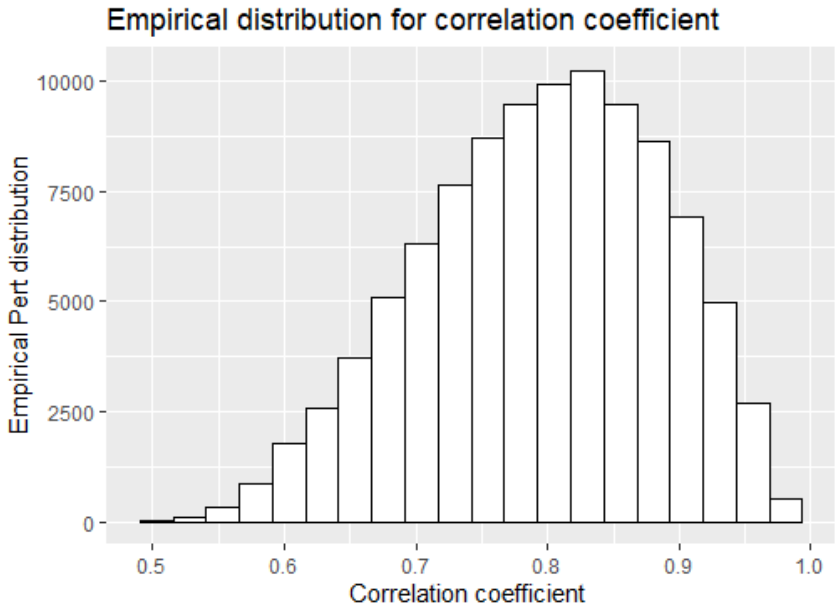


Figure C3: Uncertainty distribution for the correlation coefficient

The analyses were performed in R version R-3.6.0, Studio version 1.2.

Abbreviations

AIC	Aikake information criteria
BIC	Bayesian information criteria
BoE	Body of evidence
CI	Confidence interval
c_c	Correlation Coefficient for Cross-over Studies
c	Correlation coefficient
CV	Coefficient of variation
DRV	Dietary reference value
EFSA	European Food Safety Authority
E%	Energy intake percentage
EKE	Expert knowledge elicitation
FG	Fasting glucose
FI	Fasting insulin
LL	Lower boundary
LP	Location parameter
PI	Prediction interval
RCS	Restricted cubic spline
RCT	Randomised control trial
RefID	Reference ID of the study
REML	REstricted Maximum Likelihood
RoB	Risk of Bias
SD	Standard deviation
SE	Standard error
SSB	Sugar-sweetened beverages
UA	Uric acid
UL	Upper boundary
VAT	Visceral adipose tissue
VP	Variation parameter
VPT	Variation parameter type
WG	Working group